



UNITED STATES PATENT AND TRADEMARK OFFICE

AB

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/989,920	11/21/2001	Roberto A. Macina	DEX-0291	2846

7590 12/09/2004  
Licata & Tyrrell P.C.  
66 East Main Street  
Marlton, NJ 08053

EXAMINER

YU, MISOOK

ART UNIT PAPER NUMBER

1642

DATE MAILED: 12/09/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

## Office Action Summary

Application No. 09/989,920	Applicant(s) MACINA ET AL.
Examiner MISOOK YU, Ph.D.	Art Unit 1642

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1)  Responsive to communication(s) filed on 27 September 2004.
- 2a)  This action is **FINAL**.                      2b)  This action is non-final.
- 3)  Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4)  Claim(s) 1-17 is/are pending in the application.
- 4a) Of the above claim(s) 6, 10-14 and 16 is/are withdrawn from consideration.
- 5)  Claim(s) \_\_\_\_\_ is/are allowed.
- 6)  Claim(s) 1-5, 7-9, 15 and 17 is/are rejected.
- 7)  Claim(s) \_\_\_\_\_ is/are objected to.
- 8)  Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9)  The specification is objected to by the Examiner.
- 10)  The drawing(s) filed on \_\_\_\_\_ is/are: a)  accepted or b)  objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11)  The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12)  Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a)  All    b)  Some \*    c)  None of:
- Certified copies of the priority documents have been received.
  - Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- 1)  Notice of References Cited (PTO-892)
- 2)  Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3)  Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date 10/29/04, 09/27/02.
- 4)  Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_\_.
- 5)  Notice of Informal Patent Application (PTO-152)
- 6)  Other: Exhibits A-D.

Art Unit: 1642

**DETAILED ACTION*****Election/Restrictions***

Applicant's election with traverse of group 100 encompassing claims 1-5, 7-9, 15, and 17 partially, relating SEQ ID NO:100 nucleic acid molecule, and the nucleic acid encoding the protein encoded by the nucleic acid molecules in the reply filed on 09/27/2004 is acknowledged. The traversal is on the ground(s) that searching of the nucleic acids, polypeptide, and antibody all together would not put a serious burden on the Examiner, and the restriction groups are not independent or distinct. Applicant also argues that MPEP 803.04 says that up to 10 nucleotide sequences are examined in a single application. These arguments have been fully considered but found unpersuasive for the following reasons. First, all the nucleic acid sequences in the instant application are a large sequences, for example, the elected SEQ ID NO:100 is close to 3 kb, which requires search in sequence databases and non-patent literature. Each of the different nucleotides sequences appear to be either cDNA, or genomic DNA from the different human chromosomes without any relationship to each other. Examination of up to 10 nucleotide sequences applies when the search burden of up to 10 nucleotide sequences is not serious on the examiner or the office's available resources, for example the use of computer time. In the instant cases, each SEQ ID NO is a large sequences and search of one application up to 10 nucleotides sequences would put a serious burden on the Office.

As for the examination of nucleic acids along with the protein encoded by the nucleic acid and the antibody binding to the protein, the polypeptide groups, and

Art Unit: 1642

antibody groups are different inventions. Polypeptides, which are composed of amino acids, and polynucleotides, which are composed of purine and pyrimidine units, are structurally distinct molecules; any relationship between a polynucleotide and polypeptide is dependent upon the information provided by the nucleic acid sequence open reading frame as it corresponds to the primary amino acid sequence of the encoded polypeptide. In the present claims, a polynucleotide of group I does not necessarily encode a polypeptide. In fact, the elected SEQ ID NO:100 does not encode any protein in the instant application. Alignment of instant SEQ ID NO:100 against the protein sequences i.e. SEQ ID NOs 165-284 reveals none of the protein is encoded by instant SEQ ID NO:100. Note Exhibit A (total 20 pages). Further, the nucleic acid molecule of claim 1(c) is complementary to the coding sequence. If SEQ ID NO:100 is a coding sequence, and then the complementary sequence would not encode any polypeptide because the complimentary sequence of a human cDNA does not usually encode any protein. The instant specification does not teach any evidence to the contrary. There is search burden also in the non-patent literature. Prior to the concomitant isolation and expression of the sequence of interest there may be journal articles devoted solely to polypeptides which would not have described the polynucleotide. Similarly, there may have been "classical" genetics papers, which had no knowledge of the polypeptide but spoke to the gene. Searching, therefore is not coextensive. The polypeptide of groups and the antibody groups are patentably distinct for the following reasons: While the inventions of both groups are polypeptides, in this instance the polypeptide is a single chain molecule, whereas antibody encompasses

Art Unit: 1642

IgG which comprises 2 heavy and 2 light chains containing constant and variable regions, and including framework regions which act as a scaffold for the 6 complementarity determining regions (CDRs) that function to bind an epitope. Therefore they are structurally distinct molecules; any relationship between a polypeptide and an antibody is dependent upon the correlation between the scope of the polypeptides that the antibody binds and the scope of the antibodies that would be generated upon immunization with the polypeptide.

The polynucleotide and the antibody group are patentably distinct. An antibody for example, IgG molecule which comprise 2 heavy and 2 light chains containing constant and variable regions, and including framework regions which act as a scaffold for the 6 complementarity determining regions (CDRs). Polypeptides, such as the antibody of group II which are composed of amino acids, and polynucleotides, which are composed of nucleic acids, are structurally distinct molecules; any relationship between a polynucleotide and polypeptide is dependent upon the information provided by the nucleic acid sequence open reading frame as it corresponds to the primary amino acid sequence of the encoded polypeptide. In the present claims, a polynucleotide in the nucleic acid groups will not encode an antibody in antibody groups, and the antibody of group III cannot be encoded by a polynucleotide of group I. Therefore the antibody and polynucleotide are patentably distinct.

The requirement is still deemed proper and is therefore made FINAL.

Art Unit: 1642

Claims 6, and 10-14, and 16 are withdrawn from further consideration pursuant to 37 CFR 1.142(b), as being drawn to a nonelected invention, there being no allowable generic or linking claim.

Claims 1-17 are pending, and claims 1-5, 7-9, 15 and 17 are examined as to the extent they are drawn to the elected invention of SEQ ID NO:100 or related nucleic acid sequence.

### ***Specification***

The disclosure is objected to because it contains an embedded hyperlink and/or other form of browser-executable code, for example at pages 54, 61. Applicant is kindly requested to go through the entire specification very carefully to see if there is any other embedded hyperlink and/or other form of browser-executable code. Applicant is required to delete all embedded hyperlink and/or other form of browser-executable code(s) in the specification. See MPEP § 608.01.

### ***Claim Objections***

Claims 1-5, 7-9, 15, and 17 are objected to because of the following informalities: the claims are still drawn to multiple inventions. The claims have not been amended to reflect the election. Appropriate correction is required.

### ***Claim Rejections - 35 USC § 101***

Claims 1-5, 7-9, 15, and 17 are rejected under 35 U.S.C. 101 because the claimed invention is not supported by either a substantial utility, or a well established utility.

Art Unit: 1642

Claim 1-5, 7-9, 15, and 17 are interpreted as drawn to SEQ ID NO:100 nucleic acid, kit containing said nucleic acid, vector containing said nucleic acid, host cell containing said vector, method of producing protein using said host cells. The specification speculates that SEQ ID NO:100 might have utilities in making protein, making antibody, diagnostic and staging assays for lung cancer or non-cancerous diseases (at pages 93-103) or detecting a risk of cancer or presence of cancer (claim 15), method of identifying lung tissue (page 103), method of producing and modifying lung tissue such as making an artificial lung (at pages 104-105), pharmaceutical (page 105) in gene therapy and antisense therapy (pages 111-113).

These utilities are not considered to be substantial enough because neither the specification nor any art of record teaches what the biological activities of SEQ ID NO:100 are. The specification at page 6, lines 12-27 teaches that the disclosed nucleic acids are lung specific. An assay to tell whether one has lung i.e. lung tissue typing is not considered a substantial enough utility. The specification asserts that the differential expression of the sequence is used for lung cancer detection. However, the specification does not teach whether the claimed nucleic acid is under-expressed or overexpressed in lung cancer. The specification does not teach a relationship to any specific disease or establish any involvement SEQ ID NO:100. The specification does not teach which protein is encoded by SEQ ID NO:100, let alone substantial or specific use for it. None of the protein sequences disclosed in the instant application is encoded by instant SEQ ID NO:100. Note the Exhibit A. Making and purifying the protein encoded by SEQ ID NO:100 does not lead to a substantial use of the claimed invention

Art Unit: 1642

because neither the specification nor the art appears to know what the structure of the protein encoded by the claimed invention. In fact, Genbank Accession No. AC079988 (gi:18873965, #DI of IDS filed on 10/29/2004) teach a genomic DNA i.e. human BAC clone RP11-795C1 (from chromosome 2) having 98.7 % sequence identity to instant SEQ ID NO:100. Note the sequence alignment of instant SEQ ID NO:100 against Genbank Accession No. AC079988 (Exhibit B). Further, Genbank Accession No. AC079988 teach that this clone is from RPCI-11 human BAC library prepared from the blood of one male donor as disclosed in The Sanger Center and The Washington University Genome Sequencing Center (DC of IDS filed on 10/29/2004, 1998, Genome Research, vol. 8, pages 1097-1108. The Sanger Center and The Washington University Genome Sequencing Center teaches that the library was constructed for sequencing human genome, and it is not cDNA library. Therefore, SEQ ID NO:100 is most likely a genomic sequence, not specific for lung only but present in every cell of human body that contains chromosome 2.

The asserted utilities as hybridization probes, antisense, the various assays numerated in the instant application do not lead to substantial and credible uses of the claimed invention due to unknown functions of the protein encoded by the claimed invention. Nothing is specific to the sequences of the claimed invention for all of the various probe uses. Any nucleic acid can be used to, identify polymorphisms, map chromosomes, type a tissue, make transgenic animals or knockout animals. The specification does not have any substantial use for pharmaceutical compositions, diagnostic assay, and methods of treatment because the specification does not teach



Art Unit: 1642

what disease(s) is caused by malfunction of the claimed invention or the protein encoded by it. Since EQ ID NO:100 does not have a substantial utility, or a well established utility, a compound that binds to SEQ ID NO:100 does not have a substantial utility, or a well established utility.

In *Brenner v. Manson*, 148 U.S.P.Q. 689 (Sup. Ct., 1966), a process of producing a novel compound that was structurally analogous to other compounds which were known to possess anti-cancer activity was alleged to be useful because the compound produced thereby was potentially useful as an anti-tumor agent in the absence of evidence supporting this utility. The court expressed the opinion that all chemical compounds are "useful" to the chemical arts when this term is given its broadest interpretation. However, the court held that this broad interpretation was not the intended definition of "useful" as it appears in 35 U.S.C. § 101, which requires that an invention must have either an immediately obvious or fully disclosed "real world" utility. The instant claims are drawn to SEQ ID NO:100 which has undetermined function or biological significance. Until some actual and specific activity can be attributed to the nucleic acid or the protein encoded by the claimed invention encoding it, the claimed invention is incomplete.

Claims 1-5, 7-9, and 17 are also rejected under 35 U.S.C. 112, first paragraph. Specifically, since the claimed invention is not supported by either a substantial asserted utility or a well established utility for the reasons set forth above, one skilled in the art clearly would not know how to use the claimed invention.

Art Unit: 1642

***Claim Rejections - 35 USC § 112***

The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 1-5, 7-9, 15, and 17 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

The claims are interpreted as drawn to a genus of nucleic acid molecules with various of degrees of variations from SEQ ID NO:100, a genus of vectors containing said nucleic acid molecules, genus of host cells containing said vectors, a method of producing a genus of polypeptides using said host cells.

The applicable standard for the written description requirement can be found: MPEP 2163; University of California v. Eli Lilly, 43 USPQ2d 1398 at 1407; PTO Written Description Guidelines; Enzo Biochem Inc. v. Gen-Prove Inc., 63 USPQ2d 1609; Vas-Cath Inc. v. Mahurkar, 19USPQ2d 1111; and University of Rochester v. G.D. Searle & Co., 69 USPQ2d 1886 (CA FC 2004).

To provide adequate written description and evidence of possession of a claimed genus, the specification must provide sufficient distinguishing identifying characteristics of the genus. The factors to be considered include disclosure of complete or partial structure, physical and/or chemical properties, functional characteristics,

Art Unit: 1642

structure/function correlation, methods of making the claimed product, or any combination thereof. In this case, the only factor present in the claims is a partial structure in the form of percent identity or hybridization. There is not even identification of any associated function with the claimed genus of partial structures. Accordingly, in the absence of sufficient recitation of distinguishing identifying characteristics, the specification does not provide adequate written description of the claimed genus.

Vas-Cath Inc. v. Mahurkar, 19USPQ2d 1111, clearly states “applicant must convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession of the invention. The invention is, for purposes of the ‘written description’ inquiry, *whatever is now claimed.*” (See page 1117.) The specification does not “clearly allow persons of ordinary skill in the art to recognize that [he or she] invented what is claimed.” (See Vas-Cath at page 1116). As discussed above, the skilled artisan cannot envision the detailed chemical structure of the encompassed genus of nucleic acid molecules, given that the specification has only described SEQ ID NO: 100. Therefore, only isolated nucleic acid comprising SEQ ID NO:100, but not the full breadth of the claim meets the written description provision of 35 U.S.C. §112, first paragraph.

Further, claim 9 is drawn to method of producing a polypeptide using the claimed nucleic acid. However, the specification does not teach the structure of the polypeptide the instant SEQ ID NO:100 encodes, thus failing to provide an adequate written description.

Art Unit: 1642

Claims 1-5, 7-9, 15, and 17 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the **enablement requirement**. The claim(s) contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

The factors considered when determining if the disclosure satisfies the enablement requirement and whether any necessary experimentation is "undue" include, but are not limited to: 1) nature of the invention, 2) state of the prior art, 3) relative skill of those in the art, 4) level of predictability in the art, 5) existence of working examples, 6) breadth of claims, 7) amount of direction or guidance by the inventor, and 8) quantity of experimentation needed to make or use the invention. *In re Wands*, 858 F.2d 731, 737, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988).

This enablement rejection is made based on the interpretation of the claims as drawn to an isolated nucleic acid molecule comprising SEQ ID NO:100, a nucleic acid that selectively hybridizes, or at least 60 % sequence identity to instant SEQ ID NO:100 (for use in lung cancer detection (note pages 93-103, abstract) or use as a vaccine (note claim 17)). Further, claim 8 is drawn to a host cell comprising the vector of claim 8. therefore claim as currently construed is broadly interpreted to encompass host cells, which are not isolated and are comprised within an organism. Thus, the claim encompass host cells that have been transfected with the vector of claims 8 that are comprised within a transgenic animal, including nonhuman or human animals and animals treated using gene therapy (note page 111 of the specification).

Art Unit: 1642

The specification does not teach whether SEQ ID NO:100 is over-expressed or under expressed in lung cancer or any other lung disease, let alone a nucleic acid that selectively hybridizes, or at least 60 % sequence identity to instant SEQ ID NO:100 being over-expressed or under-expressed in any lung disease including lung cancer. The specification provides neither guidance on nor exemplification of how to correlate the data presented in the specification with the ability to use SEQ ID NO:100 for the assessment of cancer risk. In other words, the specification does not present any in vivo data to correlate either detection of the nucleic acid or absence of the nucleic acid to growth of any tumor. Tockman et al (Cancer Res., 1992, 52:2711s-2718s) teach considerations necessary in bringing a lung cancer biomarker to successful clinical application. Tockman et al teach that prior to the successful application of newly described lung cancer markers, research must validate the markers against acknowledged disease end points, establish quantitative criteria for marker presence/absence and confirm marker predictive value in prospective population trials (see abstract). Early stage markers of tumorigenicity have clear biological plausibility as markers of preclinical cancer and if validated can be used for population screening (p. 2713s, col 1). The reference further teaches that once selected, the sensitivity and specificity of the biomarker must be validated to a known (histology/cytology-confirmed) cancer outcome. The essential element of the validation of an early detection marker is the ability to test the marker on clinical material obtained from subjects monitored in advance of clinical cancer and link those marker results with subsequent histological confirmation of disease. This irrefutable link between antecedent marker and

Art Unit: 1642

subsequent acknowledged disease is the essence of a valid intermediate end point marker (p. 2714, see Biomarker Validation against Acknowledged Disease End Points). Clearly, prior to the successful application of newly described markers, markers must be validated against acknowledged disease end points and the marker predictive value must be confirmed in prospective population trials (p. 2716s, col 2). The specification provides insufficient guidance, and provides no working examples of correlating in lung cancer to either detection of SEQ ID NO:100 or to absence of SEQ ID NO:100, which would provide guidance to one skilled in the art to use the claimed invention without undue experimentation. Considering lack of examples and the limited teachings of the specification, and unpredictability in the art, it is concluded that undue experimentation would be required to practice the claimed invention. It is noted that law requires that the disclosure of an application shall inform those skilled in the art how to make the alleged discovery, not how to screen it for themselves.

As for claim 17, the claim is interpreted as drawn to DNA vaccine comprising SEQ ID NO:100 or other related nucleic acids as described in the base claim. Roos et al., (06/25/2004, Prostate, vol. 9999, pages 1-7) teach that DNA vaccine is used to generate immunity against a specific antigen in this case tumor antigen, and the consideration necessary for DNA vaccine to work is to put the DNA encoding the antigen of interest into a vector such that the antigen is expressed in vivo (note page 2, left column). The instant application does not even teach the structure of protein being encoded by the claimed nucleic acid. Note the attached Exhibit A, showing none of the numerous protein sequences in the instant application is being encoded by instant SEQ

Art Unit: 1642

ID NO:100. Since DNA vaccine requires testing of its efficacy *in vivo* (note page 3, right column of Roos et al.), and the instant specification does not even teach whether instant SEQ ID NO:100 encodes any protein, let alone any antigen that an animal can produce immune response against it, it requires undue experimentation to use the instant claimed vaccine. The specification provides insufficient guidance, and provides no working examples of how to use SEQ ID NO:100, and its related nucleic acid sequences as vaccine, which would provide guidance to one skilled in the art to use the claimed invention without undue experimentation. Considering lack of examples and the limited teachings of the specification, and unpredictability in the art, and broad breath, it is concluded that undue experimentation would be required to practice the claimed invention.

As for claim 8, the specification does not provide a sufficient amount of guidance, direction, or exemplification to enable the skilled artisan to make or use host cells that are comprised within a non-human transgenic animal. In addition, the specification does not teach provide a sufficient amount of guidance, direction, and exemplification to enable the skilled artisan to have a reasonable expectation of successfully producing host cells within a living organism, which comprise the vectors of claims 8, by gene transfer, or gene therapy. The art of gene therapy, i.e., the *in vivo* delivery genetic information to targeted cells within a body using naked DNA or viral vectors or by reintroducing *ex vivo* modified host cells into the body, is still in its infancy. Moreover, the art is highly unpredictable and its successful application has been hindered by numerous limitations, which the specification does not remedy and would preclude the

Art Unit: 1642

skilled artisan from having a reasonable expectation of successfully making and using the claimed invention without need of performing an undue amount of experimentation. The specification does not teach method of gene therapy using the instantly claimed nucleic acid. The art recognizes that gene therapy is not a trivial matter. Friedmann (Scientific American, June 1997, pages 96-101), Verma and Somia (1997, Nature, vol. 389, pages 239-242), and Rubanyi (2001, Molecular Aspects of Medicine 22, pages 113-142) all teach that gene therapy art still faces major hurdle to overcome. For example, the teachings of the specification have not overcome the problems with *in vivo* delivery and expression. Verma et al. (*Nature* 1997, **389**: 239-242) teach that the Achilles heel of gene therapy is gene delivery. Verma et al. state that the ongoing problem is the inability to deliver genes efficiently and to obtain sustained expression. Rubanyi at the abstract teaches that the prerequisite of successful gene therapy includes "therapeutically suitable genes with a proven role in pathophysiology of the disease". The instant specification fails at this first prerequisite because the specification does not teach any therapeutically suitable genes with a proven role in pathophysiology of the disease. Friedman summarizes the current state of gene therapy as "treating disease by providing needed gene remains a compelling idea, but clinical and basic researchers still have much to do before gene therapy can live up to its promise" (note the italicized headline at the top of page 96). The instant specification does not teach a single technical problem being solved for gene therapy art. In view of the preponderance of evidence establishing the state of the art, now and at the time the application was filed, and the level of unpredictability associated therewith, in the



Art Unit: 1642

absence of a disclosure of an amount of guidance, direction, and exemplification that is reasonably commensurate in scope with the claims, it appears that skilled artisan could not make and use the claimed invention with a reasonable expectation of success without having the need to perform an undue amount of experimentation. Amending claim 8 to recite "isolated" before "host cell" can obviate this ground of rejection.

### ***Claim Rejections - 35 USC § 102***

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

(e) the invention was described in a patent granted on an application for patent by another filed in the United States before the invention thereof by the applicant for patent, or on an international application by another who has fulfilled the requirements of paragraphs (1), (2), and (4) of section 371(c) of this title before the invention thereof by the applicant for patent.

Claims 1-5, 7, 8, 15, and 17 are rejected under 35 U.S.C. 102(e) as being anticipated by US Pat. 6,368,794 B1 (issued April 09, 2002, filed Jan. 15, 1999).

The claims are interpreted as drawn to an isolated nucleic acid molecule comprising a nucleic acid that selectively hybridizes or at least 60 % sequence identity to SEQ ID NO:100 (claim 1), wherein said nucleic acid molecule is a cDNA (claim 2), genomic DNA (claim 3), a mammalian nucleic acid molecule (claim 4), a human nucleic acid molecule (claim 5), in a vector (claim 7), in a host cell comprising said vector (claim 8), a kit comprising a means of for determining the presence of said nuclei acid, or said nucleic acid molecule in a vaccine (claim 17).

Art Unit: 1642

US Pat. 6,368,794 B1 teaches SEQ ID NO:3 having 63.5 % sequence identity to instant SEQ ID NO:100. Note the sequence alignment of instant SEQ ID NO:100 against SEQ ID NO:3 of US Pat. 6,368,794 B1 (Exhibit C). The patent teaches that SEQ ID NO:3 is a cDNA obtained by reverse transcribing polyA mRNA (note columns 17-20, claim 1) from a human tissue, and also teaches method of detecting cancerous condition using a kit comprising a means to detect the nucleic acid of the base claim i.e., claim 1 of the patent (note claims 5-8).

Since the specification at paragraph [0034] defines "exon" refers to a nucleic acid sequence found in genomic DNA that is bioinformatically predicted and/or experimentally confirmed to contribute a contiguous sequence to a mature mRNA transcript, SEQ ID NO:3 of US Pat. 6,368,794 B1 is a genomic DNA. Note the SEQ ID NO:3 meets the definition of genomic DNA as defined in the instant application since it is transcribe from exons. Note SEQ ID NO:3 of the a patent is a mature mRNA.

As for claims 7, and 8, the patent (at column 15 under the heading "Isolation and sequencing of cDNA Clones") teaches that the clones disclosed in the patent are in plasmid in a host cell.

As for claim As for claim 17, the preamble recitation of "vaccine" is merely suggestive of an intended use and is not given patentable weight for purposes of comparing the claim with the prior art. The claim reads on nucleic acids *per se*.

Thus, US Pat. 6,368,794 B1 anticipates claims 1-5, 7, 8, 15, and 17.

Art Unit: 1642

Claims 1, 3-5, 7, 8, 15, and 17 are rejected under 35 U.S.C. 102(a) as being anticipated by Genbank Accession No. AC079988 (gi:10800346 with public availability date of 10/14/2000 online, IDS #DI filed on 10/29/04).

The claims are interpreted as drawn to an isolated nucleic acid molecule comprising a nucleic acid that selectively hybridizes or at least 60 % sequence identity to instant SEQ ID NO:100 (claim 1), wherein said nucleic acid molecule is genomic DNA (claim 3), a mammalian nucleic acid molecule (claim 4), a human nucleic acid molecule (claim 6), in a vector (claim 7), in a host cell comprising said vector (claim 8), kit comprising a means to detect the nucleic acid of claim 1, or said nucleic acid molecule in a vaccine (claim 17).

Genbank Accession No. AC079988 teach a human chromosome 2 clone RP11-795C1, which is isolated from the human BAC library RPCI-11 according to AC079988 (gi:18873965), which contains an insert having 94.6 % sequence identity to instant SEQ ID NO:100. Note the sequence alignment of instant SEQ ID NO:100 against Genbank Accession No. AC079988 with gi:10800346 (Exhibit D). Genbank Accession No. AC079988 sequencing vector M13, and Dye-primer ET, and other means to detect the nucleic acid inserted in the clone.

As for claims 7, and 8, Voet et al., (1900, Biochemistry, John Wiley & Sons, pages 839-844) teach that a clone is in a host cell containing a nucleic acid of insert in an appropriate vector.

As for claims 15, and 17, the intended use in claim 15, and the preamble recitation in claim 17 are merely suggestive of an intended use and is not given

Art Unit: 1642

patentable weight for purposes of comparing the claim with the prior art. The claim reads on nucleic acids *per se*, and a means *per se*.

Thus, Genbank Accession No. AC079988 anticipates claims 1, 3-5, 7, 8, 15, and 17.

### **Conclusion**

Any inquiry concerning this communication or earlier communications from the examiner should be directed to MISOOK YU, Ph.D. whose telephone number is 571-272-0839. The examiner can normally be reached on 8 A.M. to 5:30 P.M., every other Friday off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Jeffrey C Siew can be reached on 571-272-0787. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

  
MISOOK YU, Ph.D.  
Examiner  
Art Unit 1642

Pages

22	56.5	1.1	162	1	US-09-989-920-199	Sequence 199, App
c 23	56.5	1.1	456	1	US-09-989-920-175	Sequence 175, App
24	55.5	1.1	101	1	US-09-989-920-184	Sequence 184, App
25	55	1.1	594	1	US-09-989-920-200	Sequence 200, App
26	55	1.1	594	1	US-09-989-920-249	Sequence 249, App
c 27	54.5	1.1	224	1	US-09-989-920-274	Sequence 274, App
28	54	1.1	100	1	US-09-989-920-264	Sequence 264, App
c 29	52.5	1.0	240	1	US-09-989-920-208	Sequence 208, App
c 30	52	1.0	130	1	US-09-989-920-208	Sequence 208, App
c 31	52	1.0	489	1	US-09-989-920-185	Sequence 185, App
c 32	51.5	1.0	610	1	US-09-989-920-212	Sequence 212, App
c 33	50.5	1.0	53	1	US-09-989-920-230	Sequence 230, App
c 34	49	1.0	101	1	US-09-989-920-184	Sequence 184, App
c 35	48	1.0	128	1	US-09-989-920-220	Sequence 220, App
c 36	48	0.9	252	1	US-09-989-920-204	Sequence 204, App
c 37	48	1.0	1533	1	US-09-989-920-259	Sequence 259, App
c 38	47.5	0.9	57	1	US-09-989-920-234	Sequence 234, App
c 39	47.5	0.9	88	1	US-09-989-920-196	Sequence 196, App
c 40	47.5	0.9	1533	1	US-09-989-920-259	Sequence 259, App
c 41	47	0.9	149	1	US-09-989-920-279	Sequence 279, App
c 42	46.5	0.9	149	1	US-09-989-920-187	Sequence 187, App
c 43	46	0.9	107	1	US-09-989-920-233	Sequence 233, App
c 44	46	0.9	183	1	US-09-989-920-245	Sequence 245, App
c 45	45.5	0.9	63	1	US-09-989-920-193	Sequence 193, App

ALIGNMENTS

RESULT 1  
 US-09-989-920-237  
 ; Sequence 237, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ IDS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 237  
 ; LENGTH: 146  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 ; US-09-989-920-237

Alignment Scores:					
Pred. No.:	1.21e-13	Length:	146		
Score:	254.50	Matches:	59		
Percent Similarity:	76.62%	Conservative:	0		
Best Local Similarity:	76.62%	Mismatches:	7		
Query Match:	5.02%	Indels:	15		
DB:	1	Gaps:	1		
us-09-989-920-100 (1-2754) x US-09-989-920-237 (1-146)					
QY 2139 ATGGGGCCAGGTCACCA-GCTACTGATGCGCCAGCCAGGTTGAGCTTCCTGGTCTC 2197					
Db 1 MetGlyAlaArgValProHisAlaAlaAspGlyProSerGlnValGluLeuProGlyVal 20					
QY 2198 CAGTCCGGATC-CCACTTGGCAGTCTCATGCTCAGATAGGTGGGCAAGTTCCTTTTGT 2256					
Db 21 GlnSerGlySerProLeuAlaAspLeuMetLeuSerAspArgTrpAspLysPhePheCys 40					
QY 2257 CACAGTGTGGTCTC-TGTCCTGAGGCCCTCATTGCTGGTGGTGGTGGTCTC--TGCTGGGAA 2313					
Db 41 HisSerAlaGlyLeuCysProGluIAserLeuLeuAlaGlyCysAlaHisAlaArgGlu 60					

GenCore version 5.1.6  
 Copyright (C) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: November 23, 2004, 08:50:59 ; Search time 16.5 Seconds  
 (without alignments)  
 4.832 Million cell updates/sec

Title: us-09-989-920-100  
 Perfect score: 5074  
 Sequence: 1 GCGAGAGCAGCTCAGCTT.....AAAAATAAAGATCGGGCC 2754

Scoring table: BLOSUM62  
 Xgapop 10.0 , Xgapext 0.5  
 Ygapop 10.0 , Ygapext 0.5  
 Fgapop 6.0 , Fgapext 7.0  
 Delop 6.0 , Deltext 7.0

Searched: 120 seqs, 14475 residues  
 Total number of hits satisfying chosen parameters: 240

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Command line parameters:  
 -MODEL=frame+n2p.model -DEV=soft -Qus-09-989-920-100 -DB=US09989920.pep  
 -SUFFIX=pto -OUT=align100 -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits  
 -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200  
 -THR SCORE=pcnt -THR MAX=100 -THR MIN=0 -ALIGN=45 -MODE=LOCAL -OUTFT=ptc  
 -NORF-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -NCFU=6 -NO XLPXY  
 -NEG SCORES=0 -LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : US09989920.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254.5	5.0	146	1	US-09-989-920-237
2	117	2.3	30	1	US-09-989-920-231
3	109	2.1	213	1	US-09-989-920-206
4	109	2.1	213	1	US-09-989-920-251
5	83.5	1.6	785	1	US-09-989-920-218
6	80	1.6	785	1	US-09-989-920-218
7	73	1.4	188	1	US-09-989-920-240
8	70.5	1.4	196	1	US-09-989-920-224
9	69	1.4	146	1	US-09-989-920-237
c 10	69	1.4	149	1	US-09-989-920-187
c 11	69	1.4	213	1	US-09-989-920-206
c 12	67	1.3	109	1	US-09-989-920-223
c 13	67	1.3	109	1	US-09-989-920-223
14	66	1.3	456	1	US-09-989-920-175
15	62	1.2	183	1	US-09-989-920-245
16	60.5	1.2	86	1	US-09-989-920-243
17	59	1.2	91	1	US-09-989-920-205
18	59	1.2	451	1	US-09-989-920-210
c 19	58.5	1.2	121	1	US-09-989-920-198
c 20	57.5	1.2	171	1	US-09-989-920-282
c 21	57.5	1.1	196	1	US-09-989-920-224

QY 2314 AAG-----CTTGGGGGCTTGC 2331  
 Db 61 LysAlaTrpAlaValAsnTyrGlyHisThrCysSerLeuCysGlyHisCys 77

RESULT 2

US-09-989-920-231  
 ; Sequence 231, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 231  
 ; LENGTH: 30  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 ; US-09-989-920-231

Alignment Scores:  
 Pred. No.: 0.000438 Length: 30  
 Score: 117.00 Matches: 26  
 Percent Similarity: 96.30% Conservative: 0  
 Best Local Similarity: 96.30% Mismatches: 1  
 Query Match: 2.31% Indels: 1  
 Db: 1 Gaps: 0

us-09-989-920-100 (1-2754) x US-09-989-920-231 (1-30)

QY 2540 ATGGTCAGTGGCTGTGACGCTCGGGAGTGG--GCCTGAGCTCAAAGTTGTC 2597  
 Db 1 MetGlyGlnValGlyValArgGlyProGlyGluValArgAlaLeuSerSerLysLeuSer 20

QY 2598 TACTGCCATGTTTTGTACT 2618  
 Db 21 TyrCysHisValPheValPro 27

RESULT 3

US-09-989-920-206  
 ; Sequence 206, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 206  
 ; LENGTH: 213  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 ; US-09-989-920-206

Alignment Scores:  
 Pred. No.: 0.00197 Length: 213  
 Score: 109.00 Matches: 73  
 Percent Similarity: 36.19% Conservative: 24  
 Best Local Similarity: 27.24% Mismatches: 81

Query Match: 2.15% Indels: 91  
 Db: 1 Gaps: 15  
 us-09-989-920-100 (1-2754) x US-09-989-920-206 (1-213)

QY 1554 GTTGATGCTGGAGAGGACTTTGACAGTGGTGGGACAGTGGCCCTCTCGCCAGCTC 1613  
 Db 9 ValAspGlnThrProArgThrLeuAlaThrMetGlyGlnArgAlaLeuProSerSerLeu 28

QY 1614 ACACCTTCTCTGGGGAGGCGCTGATCACCCTCCACCTAGTACCTTGGGACTGA 1673  
 Db 29 AlaLeuLeuSer-----ArgPro-LeuSerProProAlaAlaCysSerGlyAs 45

QY 1674 GGACCT-TTTGGCTTCTGGAGCCCTGCAAGCTCTTCCATGTGTCAGTGTCTTCC 1732  
 Db 45 pProGlyCysGly--SerGlyAlaGlyLeuPro-----SerAlaSe 58

QY 1733 TGCTACAAAGGAGACTGCTCACAGTGGCTCAGCTGGGTTTGAGGGGGCGCC 1792  
 Db 58 rAlaAlaAlaGlyTleAlaSerSerAlaValGluAlaValCysGlyAsAlaAlaProAl 78

QY 1793 GGCCCTCCATAGGTTACTCTGGGCTGAGAAATTGCAATCTGCCATTGCGAGATG 1852  
 Db 78 aCysLeu-----LeuArgThrProLeu----- 85

QY 1853 GCCTCAAAATGGAAGAGTCCACGGGAGATGGTCCGGAGTCC-----GGCTGTGC 1904  
 Db 86 -----ArgGlyLeuLeuLysProThrGlyProArgSerThrMetGluCysProR 102

QY 1905 -----CATCCAGCCCTGTGGCTTGTCCAGCTCTGTGCACCCCTGGTGT 1951  
 Db 102 oAlaLeuIleValHisProProThrGlyGlyMetAlaArgAlaAlaSer----- 119

QY 1952 TTCACTCCAGGGCAGACAGCCACTCAGTTCCTTTCTCGTGCAGTACAGTAGTGA 2011  
 Db 120 -----GlnProTrpAlaAlaAlaSerAlaThrProMetLeuSerSerLysAlaSer 136

QY 2012 TAGCAGCTGGGGCTAAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCACT 2071  
 Db 137 -----LeuCysIle-----ProThrGluAr 143

QY 2072 GATCTCCCTAAAGCAATGGGGAGGCCCCCAGCTAGCCAGTTTCAGGAAGTCACTGG 2131  
 Db 143 gProProGlnProLeuMetArgThrPro----- 153

QY 2132 AGTTAGATGGGGCCAGGGTCCCAGAGTACTGATGGCCGAGCCAGGTTGAGTTCCT 2191  
 Db 154 -----AlaAlaArg-SerHis-----TrpProIleProHisProAlaSerT 167

QY 2192 GG---TGTCCAGTCCGGATCCCACTT-----GCAGATCTCATGC 2227  
 Db 167 hrAlaCysProAlaProLeuProValValLeuValAlaProArgSerThrIleLeuSerM 187

QY 2228 TCTCAGATAGTGG-----GACAAGTTCT 2251  
 Db 187 etSerArgThrTrpThrCysArgArgTrpAlaValAlaProCysArgAlaGluLysLeuM 207

QY 2252 TTTCTCACAGTGGTGGCTCT 2271  
 Db 207 etCysSerSerSerArgSer 213

RESULT 4  
 US-09-989-920-251  
 ; Sequence 251, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920

; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 251  
 ; LENGTH: 213  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-251

Alignment Scores:  
 Pred. No.: 0.00197 Length: 213  
 Score: 109.00 Matches: 73  
 Percent Similarity: 36.19% Conservative: 24  
 Best Local Similarity: 27.24% Mismatches: 81  
 Query Match: 2.15% Indels: 91  
 DB: 1 Gaps: 15

us-09-989-920-100 (1-2754) x US-09-989-920-251 (1-213)

QY 1554 GTTATGCGTGGAGAGACTTTGACGAGTGGTGGGAGAGTGGCCCTGCGCCAGCTC 1613  
 Db 9 ValAspGlnThrProArgThrLeuAlaThrMetGlyGlnArgAlaLeuProSerSerLeu 28  
 QY 1614 ACACCTTGTCTGGCAGGGCAGCTGATCTCACCTCCACCTAGTACCTGGGGACTGA 1673  
 Db 29 AlaLeuLeuSer-----ArgPro-LeuSerProProAlaAlaCysSerGlyAs 45  
 QY 1674 GGACCT-TTTGGCTTCTGGAGCCTCAAGCCTTCCCATGTGTCACCTCTTCC 1732  
 Db 45 pProGlyCysGly---SerGlyAlaGlyLeuPro-----SerAlase 58  
 QY 1733 TGCTACAAAGGGGACTGCTCACAGTGGCTCAGCTGGTGGTGGTGGGGCCGCC 1792  
 Db 58 rAlaAlaAlaGlyIleAlaSerSerAlaValGluAlaValCysGlyAspAlaAlaProAl 78  
 QY 1793 GGCCTCCATAGGATATCTGGGGCTGAGAAATCTGCATCTGCCATGTGAGATGACA 1852  
 Db 78 aCysLeu-----LeuArgThrProLeu----- 85  
 QY 1853 GCCTCAATGGAAGAGTCCACCGGAGATGGTCCGAGTCC-----GGCTGTGCG-- 1904  
 Db 86 -----ArgGlyLeuLeuLeuLeuProThrGlyProArgSerThrMetGluCysProPr 102  
 QY 1905 -----CATCCAGCCCTGTGGCTTGTCCAGCCTCTGTGTCACCCCTGTGTC 1951  
 Db 102 oAlaLeuIleValHisProProThrGlyMetAlaArgAlaAlaSer----- 119  
 QY 1952 TTCCTCCAGGGGCGAGCAGCAGCAGCAGTGCAGTTCCTTCTGAGTAAACAGTAGTGA 2011  
 Db 120 -----GlnProTrpAlaAlaAlaSerAlaThrProMetLeuSerSerLysAlaSer-- 136  
 QY 2012 TAGCAGCTGGGCTPACAGGCTAGGCTTGTGTTCTCGCGATTTGGTCCAGTCTACTC 2071  
 Db 137 -----LeuCysIle-----ProThrGluAr 143  
 QY 2072 GATCTCCCTAAAGCAATGGGAGCCCGCCACTAGCCAGTTCAGGAAGTCAACTGGG 2131  
 Db 143 gProProGlnProLeuMetArgThrPro----- 153  
 QY 2132 AGTTAGATGGGGCGAGGTCACAGCTACTGATGCGCGAGCCAGGTTGAGCTCTCT 2191  
 Db 154 -----AlaAlaArg-SerHis-----TrpProIleProHisProAlaSerT 167  
 QY 2192 GG---TGTCCAGTCGGATCCCACTT-----GCAGATCTCATGC 2227  
 Db 167 hrAlaCysProAlaProLeuProValValLeuValAlaProArgSerThrIleLeuSerM 187  
 QY 2228 TCTCAGATAGTGG-----GACAGTCTCT 2251  
 Db 187 etSerArgThrTrpThrCysArgArgTrpAlaValAlaProCysArgAlaGluLysLeuM 207

QY 2252 TTTGTCCACAGTCTGGCTCT 2271  
 Db 207 etCysSerSerSerArgSer 213

RESULT 5

US-09-989-920-218  
 ; Sequence 218, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote  
 ; FILE REFERENCE: DEK-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 218  
 ; LENGTH: 785  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-218

Alignment Scores:  
 Pred. No.: 0.139 Length: 785  
 Score: 83.50 Matches: 52  
 Percent Similarity: 33.16% Conservative: 11  
 Best Local Similarity: 27.37% Mismatches: 63  
 Query Match: 1.65% Indels: 64  
 DB: 1 Gaps: 11

us-09-989-920-100 (1-2754) x US-09-989-920-218 (1-785)

QY 1463 ACAGCCCTTGTCTTCTTCCCTTCCCGCCCTTCCCTGTCCTGGGACCCGCTGGCCCTT 1522  
 Db 655 ArgThrLeuValTyrPro-----LeuGluLeuProLysGluPro----- 667  
 QY 1523 TGGTCTGCATCCCTGCCAGGTCCTCAGGGTGTGATGGGTGGAGAGGACTTTGAGCAG 1582  
 Db 668 -----ThrSerProProPheArgPro-----CysProGluProAsp---GluLys 681  
 QY 1583 TGGTGGCAGCAGTGGCTCTCGCCAGCTCACACTTGTCTGGAGGGGCGAGCCTGA 1642  
 Db 682 LeuTrpAspProValGlyTyrTyrSerAsp----- 692  
 QY 1643 TCTCACCTCCACCTAGTACCTTGGGACTGAGCACCTTTTGGCTTCTCTGGAGCCGCAA 1702  
 Db 693 GlySerLeuLysIleValProGlyHisAlaArg----- 703  
 QY 1703 GCCTTCCCATGTCTCCAGCTCTTCTCTGTACAAAGGGGACTGCTCACAGTGGCCT 1762  
 Db 704 -----CysGlnProGly----- 707  
 QY 1763 CAGCTTGGTGGTGGTGGGCGCGCCCGCCCGCCCTCCATAAGGGGTATCTTGGGCTGAG 1822  
 Db 708 -----GlyGlyProProSerProProGlyIlePro---GlyGlnProLeuProSer 724  
 QY 1823 AATTCTCACTGCCATGGAGGATGGACAGCCTCAATGGA-----AGGAGT 1870  
 Db 725 ProThrArgLeuHisLeuGlyGlyArgAsnSerAsnAlaAsnGlyTyrValArgLeu 744  
 QY 1871 CCCAGGGAGATGGTCCGAGGTCGGCTGTGGCCATCCAGCCCGCCCTGTGGCTTGTCCAG 1930  
 Db 745 GlnLeuGlyGlyGluAsp-ArgGlyGlyLeuGlyHisProLeuPro-----G1 760  
 QY 1931 CCTGTGTGCACCCCTGTCTTCCACTCCAGGGGCGAGCAGCAGCCACTGCTGCTTCTTT 1990  
 Db 760 lLeuAlaAspGluLeuArgArgLysLeuGlnArgGln-----ProLe 775







QY 2634 AAATATGCTTTAATTCAGGTACAAAACATGGCAGTAGGACAACTTTCAGCTCAACGCC 2575  
 Db 101 -----ProThrPro 103  
 QY 2574 ACCTCCAGGACCTCTCACCCACCTG-ACCCATCCAAGGGCCACACCA---CCCCGA 2519  
 Db 104 ThrThrProHisProProThrProHisProProHisProGlnHisProHisProArg 123  
 QY 2518 CAGATACTCCACCACTTAGAAAAGATCACCA 2483  
 Db 124 -----ProProProThrSerThrHisPro 131

RESULT 10  
 US-09-989-920-187  
 ; Sequence 187, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-yu  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Sun, Yongming  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; PRIOR FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 187  
 ; LENGTH: 149  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-187

Alignment Scores:  
 Pred. No.: 1-21 Length: 149  
 Score: 69.00 Matches: 32  
 Percent Similarity: 37.04% Conservative: 8  
 Best Local Similarity: 29.63% Mismatches: 44  
 Query Match: 1.39% Indels: 24  
 DB: 1 Gaps: 4

us-09-989-920-100 (1-2754) x US-09-989-920-187 (1-149)  
 QY 329 GGAGTGGCAATCAGAAAGGTTACAAACTTCTACCTGCAACCCCTGGCCGTTAAGA 270  
 Db 25 GlyGlyGlyGlnGlnlys-----PheArgValAspMetProGly----- 37  
 QY 269 TGACACAGCATGCAACCCCGAGGAGTGTGTGAAGCCAGTCAATGTTGGAGACC 210  
 Db 38 -----SerGlySerAlaPheIleProThrIleAsnAlaIleThr 50  
 QY 209 CGCTTAGACCTTCCCTGCTGGGACATCGCCAGTGCATTTACACTGTCTAGA 150  
 Db 51 ThrSerGlnAspLeuGln-----TrpMetValGlnProThrValIleThrSerMetSer 68  
 QY 149 AAAGCTAAGCCACCACTCCACTCTCCAGGAGGTACACAAAGATCAAAAGAGGAT 90  
 Db 69 AspProTyProArgSerHisProTySerProLeuProGlyLeuAlaSerValAlaGly 88  
 QY 89 CATCGAGCT-----ACCAGGAGCCTGTTTCCCAATG 54  
 Db 89 HisMetAlaLeuProArgProGlyValIleLysThrIleGlyThrValGlyArgArg 108  
 QY 53 CGAAGGGAACAGCATCTCCA 30  
 Db 109 ArgArgAspGlnLeuSerPro 116

RESULT 11  
 US-09-989-920-206  
 ; Sequence 206, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-yu  
 ; APPLICANT: Sun, Yongming

; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-yu  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Sun, Yongming  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot.  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 206  
 ; LENGTH: 213  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-206

Alignment Scores:  
 Pred. No.: 1-25 Length: 213  
 Score: 69.00 Matches: 48  
 Percent Similarity: 39.31% Conservative: 20  
 Best Local Similarity: 27.75% Mismatches: 67  
 Query Match: 1.39% Indels: 40  
 DB: 1 Gaps: 10

us-09-989-920-100 (1-2754) x US-09-989-920-206 (1-213)  
 QY 658 CCATCAGCTCCCTGCACAGGACGAGAAACCCATCCCTCAGAGCTACTTCCCTAAGGACA 599  
 Db 13 ProArgThrLeuAlaThrMetGlyGlnArgAlaLeuProSerSerLeuAlaLeuLeuSer 32  
 QY 598 CTGCTTAGAAGGCCCTTAAGTGGCA-----GAGCCGGGATG---GCAGGTGCA 551  
 Db 33 ArgProLeuSerProProAlaAlaCysSerGlyAspProGlyCysGlySerGlyAla 52  
 QY 550 GCCTTGGCCACTAGCACAGGATGGCC---CTCAGATCTGTAGCTGTGACCCATAGATT 494  
 Db 53 GlyLeuProSerAlaSerAlaAlaGlyIleAlaSerSerAlaValGluAlaValCys 72  
 QY 493 GGGGAGGGAGAAAGCC-----AGGAGAGCTTTTGCAGAAAGGCTCGGAAGCTAAGT 440  
 Db 73 GlyAspAlaAlaProAlaCysLeuLeuArgThrProLeuArgGlyLeuLeuLeuProThr 92  
 QY 439 GGGGTTTTCAGGAGCTTCAGGG-TGCCCTGGGCGAGTCCCAAGGAGGCGCTGGGGTGCC 381  
 Db 93 GlyPro--ArgSerThrMetGluCysPro----- 101  
 QY 380 TGACCCGCGCTCA-CTGGCCCTCAGCCCTGTCCCTAAGGGCCGATGGGAAGAGGTGG 322  
 Db 102 -----ProAlaLeuIleValHisPro---ProThrGlyGlyMetAlaArgAla 117  
 QY 321 CAATCAGAGAGGGTACAAACATTTCTACCTGCAACCCCTGGACCCGTTAAGATGACACAG 262  
 Db 118 AlaSer-----GlnProTrpAlaAlaAlaSerAlaThr 128  
 QY 261 AGCATGCAACCCAGCGGAAAGTCCCTGTGTGAAGCCAGTCAATGTGGAGACCCGCTCTAG 202  
 Db 129 ProMetLeuSerSerLysAlaSerLeuCys-----IleProThrGlu 142  
 QY 201 AGACTTGGCCCT---GCTCTGGGACATCGCCAGTCA 166  
 Db 143 ArgProProGlnProLeuMetArgThrProAlaAla 155

RESULT 12  
 US-09-989-920-251  
 ; Sequence 251, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-yu  
 ; APPLICANT: Sun, Yongming

/ APPLICANT: Liu, Chenghua  
 / TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins  
 / FILE REFERENCE: DEX-0291  
 / CURRENT APPLICATION NUMBER: US/09/989,920  
 / CURRENT FILING DATE: 2001-11-21  
 / PRIOR APPLICATION NUMBER: 60/252,500  
 / PRIOR FILING DATE: 2000-11-22  
 / NUMBER OF SEQ ID NOS: 284  
 / SOFTWARE: Patent in version 3.1  
 / SEQ ID NO 251  
 / LENGTH: 213  
 / TYPE: PRT  
 / ORGANISM: Homo sapien  
 US-09-989-920-251

Alignment Scores:  
 Pred. No.: 1.25 Length: 213  
 Score: 59.00 Matches: 48  
 Percent Similarity: 39.31% Conservatve: 20  
 Best Local Similarity: 27.75% Mismatches: 67  
 Query Match: 1.39% Indels: 40  
 DB: 1 Gaps: 10

us-09-989-920-100 (1-2754) x US-09-989-920-251 (1-213)  
 QY 658 CCATCAGTCCCTGCACAGAGCAGAAAACCCATGCTCAGAGTACTTCTTAAGGACA 599  
 Db 13 ProArgThrLeuAlaThrMetGlyGlnArgAlaLeuProSerSerLeuAlaLeuLeuSer 32  
 QY 598 CTGGCTTAGAAGCCCTAANGTGC-----CAGCCGGGATG---GCAGTGCA 551  
 Db 33 ArgProLeuSerProProAlaAlaCysSerGlyAspProGlyCysGlySerGlyAla 52  
 QY 550 GCCCTGGCCACTACACAGCAGTGTCC---CTCAGATCTGTAGCTGTGACCCATAGATT 494  
 Db 53 GlyLeuProSerAlaSerAlaAlaAlaGlyLeAlaSerSerAlaValGluAlaValCys 72  
 QY 493 GGGGGAGGAGAAAGCC-----AGGAGAGCTTTTCAGAAAGTCTCGAAGCTAAGT 440  
 Db 73 GlyAspAlaAlaProAlaCysLeuLeuArgThrProLeuArgGlyLeuLeuLysProThr 92  
 QY 439 GGGGTTTCAGGACTCTCAGG-TGCCCTGGCAGGTCCCAAGGAGGCTGGGGTCC 381  
 Db 93 GlyPro---ArgSerThrMetGluCysPro----- 101  
 QY 380 TGACCAGGCTCA-CTGGCCCTCAAGCCCTGCTCCTAAGGGCGATGGGAAGGATGG 322  
 Db 102 -----ProAlaLeuIleValHisPro---ProThrGlyGlyMetAlaArgAla 117  
 QY 321 CAATCAGAGGGGTACAAACATTTCTACCTGCAACCCTGGACCGGTAAAGTACACAG 262  
 Db 118 AlaSer-----GlnProTTPAlaAlaAlaSerAlaThr 128  
 QY 261 AGCATGCAACCCGAGCGGAAGTGCCTGTGTGAAGCCAGTCAATGGGAGACCCCTCTAG 202  
 Db 129 ProMetLeuSerSerLysAlaSerLeuLys-----IleProThrGlu 142  
 QY 201 AGACCTTCCCT---GCTCTGGGACATCGCCAGTGCA 166  
 Db 143 ArgProProGlnProLeuMetArgThrProAlaAla 155

RESULT 13  
 US-09-989-920-223  
 / Sequence 223, Application US/09989920  
 / GENERAL INFORMATION:  
 / APPLICANT: Macina, Roberto  
 / APPLICANT: Recipon, Herve  
 / APPLICANT: Chen, Sei-Yu  
 / APPLICANT: Sun, Yongming  
 / APPLICANT: Liu, Chenghua  
 / TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins  
 / FILE REFERENCE: DEX-0291  
 / CURRENT APPLICATION NUMBER: US/09/989,920

/ CURRENT FILING DATE: 2001-11-21  
 / PRIOR APPLICATION NUMBER: 60/252,500  
 / PRIOR FILING DATE: 2000-11-22  
 / NUMBER OF SEQ ID NOS: 284  
 / SOFTWARE: Patent in version 3.1  
 / SEQ ID NO 223  
 / LENGTH: 109  
 / TYPE: PRT  
 / ORGANISM: Homo sapien  
 US-09-989-920-223

Alignment Scores:  
 Pred. No.: 1.61 Length: 109  
 Score: 67.00 Matches: 34  
 Percent Similarity: 28.87% Conservatve: 7  
 Best Local Similarity: 23.94% Mismatches: 41  
 Query Match: 1.32% Indels: 61  
 DB: 1 Gaps: 8

us-09-989-920-100 (1-2754) x US-09-989-920-223 (1-109)  
 QY 2294 TGGGTGCTCTGCTGGGAAAAGCTTTCCGGGGTTCGTTGTTAACCCACAGAGAAAG 2353  
 Db 3 TrpLeuSerValGlyGly-----GlyGlyArgGluTrp-----SerGluMetLeu 17  
 QY 2354 GGCACCTTTGGGTGGCCCTCTTCGACGCTCCCGTGTGGTGGAGCAGCGTTACTGT 2413  
 Db 18 GlyValValTrp-----TTPTrpGlyGlyValGly 27  
 QY 2414 GTTCTC-----TAATGTTTATGTTTAAATAAGTATTTCTTAAAGATGTA 2461  
 Db 28 ValTrpValGlyValGlyValCysGlyVal----- 38  
 QY 2462 CCTCCACACCTTTCACAGATGGGTGACTCTTTTCTAAAGTGGTGGGATCTGTGCG 2521  
 Db 39 -----TrpTrpValValGlyValTrpTrpArgCysValGly 52  
 QY 2522 GGGTGGTGGCCCTGGATG---GGTCAGGTGGGTGTGAGAGGTCC-TGGGGAGGTGG 2577  
 Db 53 CysGlyCysValValTrpGlyGlyValGlyValGlyValGlyCysTrpGlyGly--- 71  
 QY 2578 CGFTGAGCTCAAAGTTGCTTACTCCCAATGTTTTGTACTCCCAATGTTTTGTA 2637  
 Db 71 ----- 71  
 QY 2638 ACTTGTACTGTACCATA-----GTGCGGACGAGAGTCTGTATGTGGGATCTGTG 2688  
 Db 72 ---CysValCysValValGlyValCysValCysValGlyGlyValGlyValGlyVal 90  
 QY 2689 CTGGG 2694  
 Db 91 ValGly 92

RESULT 14  
 US-09-989-920-175  
 / Sequence 175, Application US/09989920  
 / GENERAL INFORMATION:  
 / APPLICANT: Macina, Roberto  
 / APPLICANT: Recipon, Herve  
 / APPLICANT: Chen, Sei-Yu  
 / APPLICANT: Sun, Yongming  
 / APPLICANT: Liu, Chenghua  
 / TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins  
 / FILE REFERENCE: DEX-0291  
 / CURRENT APPLICATION NUMBER: US/09/989,920  
 / CURRENT FILING DATE: 2001-11-21  
 / PRIOR APPLICATION NUMBER: 60/252,500  
 / NUMBER OF SEQ ID NOS: 284  
 / SOFTWARE: Patent in version 3.1  
 / SEQ ID NO 175  
 / LENGTH: 456  
 / TYPE: PRT

ORGANISM: Homo sapien  
 US-09-989-920-175

Alignment Scores:  
 Pred. No.: 2.17 Length: 456  
 Score: 66.00 Matches: 115  
 Percent Similarity: 32.51% Conservative: 56  
 Best Local Similarity: 21.86% Mismatch: 184  
 Query Match: 1.30% Indels: 174  
 Gaps: 1

US-09-989-920-100 (1-2754) x US-09-989-920-175 (1-456)

QY 885 CCGTGTGTGACTGCCCTGTGACAGTAGCAGTAGCGGTGGGCCACCTGGCCCACTGGCTGT 944  
 Db 3 ProArgCysProProCys-----GlyCys 11

QY 945 TTATGCCACTGATTATG-----ATAGGAATATTACTTTGAA----- 983  
 Db 12 ---CysGluArgLeuValLeuAsnValalaglyLeuArgPheGluThrArgAlaArgThr 30

QY 984 -----CCCAATGAAGTGTTCCTCCCATCACAAAATAAATAATTTCTTA--- 1028  
 Db 31 LeuGlyArgPheProAspThrLeuLeuGlyAspProAlaArgArgGlyArgPheTyrAsp 50

QY 1029 -----TTTTTAGTACATGATTATACCAAAAATATGACTCAATTATTG 1073  
 Db 51 AspAlaArgGluTyrPhePheAspArgHisArgProSerPheAspAlaValLeu--- 69

QY 1074 TATTTTGGATTTTATCAATTTAAATAATGTGAAATTTGTTT-----GCTCTACGGCCA 1127  
 Db 70 -----Tyr-Tyr-TyrGlnSerGlyGlyArgLeuArgArgProAlaHisValPro 85

QY 1128 ACATATATTGATTTGCTCTGCTGCTGAAAGCCCAAAATATTACCGTACGCCGT 1187  
 Db 86 LeuAspValPheLeuGluValAlaPheTyrGlyLeuGlyAlaAlaLeuAlaArg 105

QY 1188 TAGAGAAAAGTCTGCTACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTGGATT 1247  
 Db 106 LeuArgGluAsp-----GluGlyCysProVal 114

QY 1248 ATTTAAGAAGCCTCAGACAGTAGGGCTTTTAAAGAATAAATGACTTGT-TTG 1306  
 Db 115 ProProGluArgProLeuProArgArgAlaPheAlaArgGln-----LeuTrpLeu 131

QY 1307 CGTTTGAAGCAGGGAAGCATTACAGTAGCG-----GTTTCTGCATTA 1351  
 Db 132 LeuPheGluPheProGluSerSerGlnAlaAlaArgValLeuAlaValSerValLeu 151

QY 1352 ACCCTGCCTATACCGCATCTGCTCTGCTGCTGGTGGCGAGCCCTTG----- 1402  
 Db 152 ValLeuLeuValSerIleValPheCysLeuGluThrLeuProAspPheArgAspAsp 171

QY 1403 ---GAAGTTCCTGGT---GCTTCAGTGGCTCCTCGAGAGTCCACCCCGCTCGTGGTGG 1456  
 Db 172 ArgAspGlyThrGlyLeuAlaAlaAlaAlaAlaAlaGlyProPheProAla----- 188

QY 1457 GAATGCAGACCCCTTTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1516  
 Db 189 -----ProLeuAsnGlySerSerGlnMetProGlyAsnProPro 201

QY 1517 GGCCTTTGGTCTGCATCCCTGCCAGTCCCTCAGGTTGTATGCTGGAGGAGCATT 1576  
 Db 202 ArgLeuProPheAsnAspProPhePheValValGluThrLeuCysIle----- 217

QY 1577 GAGCAGTGGTGGG-AGCAGTGGCTCTCGCCAGCTCACACTTCTGCTGGAGGGGC 1635  
 Db 218 -----CysTrpPheSerPheGluLeuValValArgLeuLeuValCysProSerLysAla 235

QY 1636 AGCCTGATCACCCTCCACCTAGTACCTTGGGACTGGAGCCTTTGCTCTCTGGAG 1695  
 Db 236 IlePhePheLys-----AsnValMetAsnLeuIleAspPheValAlaIleLeu--- 251

1696 CCTGCAAGCCTTCCCAATGTGTCCA-GCTGCTTCTTCTGTCTACAAAGGGGACTGCTCAC 1754  
 252 -----ProTyrPheValAlaLeuGlyThrGluLeuAlaAaArgGlnArgGlyValGlyGln 269

1755 AGTGGCTCAGCTTGTGTTTGGG-----GGCCGCCCCCGC 1793  
 270 GlnAlaMetSerLeuAlaIleLeuArgValIleArgLeuValArgValPheArgIlePhe 289

1794 GCCCTCATAGGGTATCTCGGCTGAGAAATTCGCATCTCCCATTCGAGATGGACAG 1853  
 290 LysLeuSerArgHisSerLysGlyLeuGlnIleLeu-----GlyGln 303

1854 CCTCAATGAAAGGAGTCCACGGGAGATGGT----- 1886  
 304 ---ThrLeuArgAlaSerMetArgGluLeuGlyLeuLeuIlePheLeuPheIleGly 322

1886 ----- 1886

323 ValValLeuPheSerSerAlaValTyrPheAlaGluValAspArgValAspSerHisPhe 342

1887 -----CCGAGGTCC-----GGCTGTGGC 1904

343 ThrSerIleProGluSerPheTrpAlaValValThrMetThrValGlyTyrGly 362

1905 CATCCAGCCCTGTGGC-----TTGTCAGGCTGTGTCACCCCTGGTG 1949

363 AspMetAlaProValThrValGlyGlyLysIleValGlySerLeuCysAla----- 379

1950 TCTTCACTCCAGGG-----CAGACAGCAGCCACTGCAGTTCCTTCTCGTAGTAAACAGT 2006

380 -----IleAlaGlyValLeuThrIleSerLeuProValIleValSerAsnPhe 397

2007 AGT-----GATAGCAGCTGGGCTAACAGCTAGGCTTGTCTGCTGCCCAT 2054

398 SerTyrPheTyrHisArgGluThrGluGlyGluAlaGlyMet----- 412

2055 TGCTCAGTCTCCTCAGTCCCTCCCTAAAGCAATGGGAGCCCGCCACTAGCCAGTTT 2114

413 -----PheSerHis-----ValAspMetGlnProCysGlyProLeu 424

2115 TCAGGAAGTCAACTGGGAGGTTAGATGGGGCCAGGGTCCACAGCTA-----CTG 2165

425 GluGlyLysAlaAsnGlyLeuValAspGlyGluValProGluLeuProProLeu 444

2166 ATGCCCGCCAGGCT 2181  
 445 -TrpAlaProProGly 449

RESULT 15  
 US-09-989-920-245  
 ; Sequence 245, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 245  
 ; LENGTH: 183  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 ; US-09-989-920-245

Alignment Scores: 3.77 Length: 183  
 Pred. No.: 183

Score: 62.00 Matches: 51  
 Percent Similarity: 29.15% Conservative: 14  
 Best Local Similarity: 22.87% Mismatches: 61  
 Query Match: 1.22% Indels: 97  
 DB: 1 Gaps: 14

us-09-989-920-100 (1-2754) x US-09-989-920-245 (1-183)

QY 1396 CCCCTGGAGGTTCTGGGCTTCAGCTGGCTCCCTGACAGT---CCACCCCGCCTC--- 1449  
 Db 3 ProAlaGlyValProTrpCysHisLeuGlySerLeuGlnProLeuProProArgPheLys 22  
 QY 1450 ---GTGGTGGGAATGAGAGCCCTTTGCTTCTTCCCTCCCGCTTCCTGTCCTGG 1506  
 Db 23 AlaValPheSerArgLeuAlaProSerLeuGluTyr---AlaTrp 36  
 QY 1507 GGAACCCGCTGGGCT--- 1521  
 Db 37 AspTyrArgAlaProThrSerHisAlaArgLeuLeuSerLeuAlaPheLeuValGluThr 56  
 QY 1522 -----TTGGTGTGCATCCCTGGCCAGGT----- 1545  
 Db 57 GlyPheSerProThrValAlaArgLeuValSerAsnSerTrpProProValValArgPro 76  
 QY 1546 -----CCCTCAGGTTGATCGCTGGAGAGGACTTGACAGTGTGGCGAGCTGGC 1599  
 Db 77 ProLeuProSer-----GlnSerAlaGly 84  
 QY 1600 CTCCTGCCAGCTCACACTCTTGTCTCTGGAGGGGCGAGCTGATCTCACCTCCACTAGT 1659  
 Db 85 IleThrGlyValGlyProProCysLeuAlaArgProIleLeu-----ProProHis 101  
 QY 1660 ACCTTGGGACTGAGACCTTTGGCTTCTCTGGAGCCCTCAAGCCTTCCCATGTGC 1719  
 Db 102 -Pro-----PhePhePhePheAspMetGlu-----SerHisAlaIle 114  
 QY 1720 CAGTGTCTTCTCTGCTACAAAGGGGACTGCTCACAGTGGCCCTCAGCTGGTGGTTTGA 1779  
 Db 114 eThrGlnAlaGlyVal-----GlnTrpArgHisLeuGlySerLeuGly 128  
 QY 1780 GGGCGCGCCCC----- 1792  
 Db 128 nProProProProMetPheLysAlaSerSerCysLeuSerLeuLeuSerSerTrpAspTyr 148  
 QY 1793 ----GGCCCTCAATAGGATCTCTGGGCTGAGAAATCTGCATCTGCCATTCGAGGATG 1848  
 Db 148 rArgArgPro-----ProArgPro-----AlaIlePheCysIle 160  
 QY 1849 GACAGCTCAAATGAGAGGACTCCACGGGAGATGGTCCGAGGCTGGCCATC 1908  
 Db 160 ePheSerArgAspGlyValSerProCysAlaProGly-----TrpSerAr 175  
 QY 1909 CAGCCCC 1915  
 Db 175 gSerPro 177

RESULT 16  
 US-09-989-920-243  
 ; Sequence 243, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Hervé  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 243  
 ; LENGTH: 86  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 ; US-09-989-920-243  
 Alignment Scores:  
 Pred. No.: 4.44 Length: 86  
 Score: 60.50 Matches: 21  
 Percent Similarity: 33.33% Conservative: 9  
 Best Local Similarity: 23.33% Mismatches: 19  
 Query Match: 1.19% Indels: 41  
 DB: 1 Gaps: 5  
 us-09-989-920-100 (1-2754) x US-09-989-920-243 (1-86)  
 QY 2321 CGGGGCTTGGTTGGTTAACCCACAGAAAGAGGGGACTGTTGGGGTGCCTCTCTCAG 2380  
 Db 9 AlaGlyValGlnTrpCysAsnPro-----GlySerLeuGln 20  
 QY 2381 CCTCCCGTGGTGGTGAAGCAGCGTACTGTGTTCTCTAATGTTCTATTTAAAT 2440  
 Db 21 ProProProProGlyPheLeuLysArgPheCysLeu-----CysLeuProSer 36  
 QY 2441 GATTTCTTTAAAGATGTAAACCTCCACAC----- 2470  
 Db 37 SerTrpGlyTyrArgHisThrProProArgProAlaAsnPheCysValPheGlyArgAsp 56  
 QY 2471 ---CTTCTCCAGATGGGTGACTTCTTCTAAAGGTGGTGGAGTATCTCTCGGGTGG 2527  
 Db 57 GlyValSerProCysTrp-----ProGlyTrp 65  
 QY 2528 -----TGTGGCCCT 2536  
 Db 66 SerLeuSerLeuAspValIleCysAspPro 75  
 RESULT 17  
 US-09-989-920-205  
 ; Sequence 205, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Hervé  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 205  
 ; LENGTH: 91  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 ; US-09-989-920-205  
 Alignment Scores:  
 Pred. No.: 5.68 Length: 91  
 Score: 59.00 Matches: 21  
 Percent Similarity: 33.80% Conservative: 3  
 Best Local Similarity: 29.58% Mismatches: 19  
 Query Match: 1.16% Indels: 28  
 DB: 1 Gaps: 4  
 us-09-989-920-100 (1-2754) x US-09-989-920-205 (1-91)  
 QY 1353 CCCTGC-----CTATCACGCATCTCGTCTCGTGTGCTGGCGAGCC----- 1397  
 Db 28 ProCysGlyArgGlyLeuAspLeuLysGlnCysProLeuTrpLeuLeuProTrpLeu 47



QY 1516 CAGCCGGTCCCA-----GGAACAGAAAGCAGGGGGCAAGAGGAAA 1475  
 Db 59 chiSaenSerProProLysTyArgGluThrThrGlyArgThrThrGlnHisThrLy 79  
 QY 1474 GAAAGGGCTGCAATCCACACAGAGCG-----GGGTGACTCT 1433  
 Db 79 sArgHisAsnThrGlnAsnHisHisThrAlaThrProAlaHisArgGlnArgThrAr 99  
 QY 1432 GAGGAGCCAGTGAAGCACCAGACCTTCCAGGGGGCTCCAGCCACACAGGACAC 1373  
 Db 99 gArgGluGlnLysGluLysGlyGlnGlnLysLysAlaSerSerThrIleThrGlnSe 119  
 QY 1372 G-----AGATGGCTATAGCAGCGGTAAATGCAGAACCCGCTCAT----- 1333  
 Db 119 xHisAspLysLysArgArgThrMetThrLysThrSerSerThrArgHisArgGlnAs 139  
 QY 1332 -CTGAAGTCCCTCCCTGCTCCAGCGCAACCAAGTCATTTTATCTTTAAAAAGCC 1274  
 Db 139 pLysSerLysLysAspArgThrArgGlnLysThrThrArgAspGluThrThrLysLysPr 159  
 QY 1273 C 1273  
 Db 159 o 159

RESULT 21  
 US-09-989-920-224  
 ; Sequence 224, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Hervé  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 224  
 ; LENGTH: 196  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-224

Alignment Scores:  
 Pred. No.: 7.73 Length: 196  
 Score: 57.50 Matches: 40  
 Percent Similarity: 24.74% Conservative: 7  
 Best Local Similarity: 21.05% Mismatches: 56  
 Query Match: 1.13% Indels: 78  
 DB: 1 Gaps: 7

us-09-989-920-100 (1-2754) x US-09-989-920-224 (1-196)  
 QY 386 CCAGGCTCTGGGACTGCCAGGGCCCTCCAGGAGCTGCTGAAACCCCC----- 439  
 Db 72 ProAlaProProProHisThrProAlaProProThrThrArgProThrThrProProPro 91  
 QY 440 -----ACTTAGCTCCAGACCTTTCTGCAAAAGCTCTCTGCTTCTCCCTCCCTCC 493  
 Db 92 ThrHisThrHisThrProThrThrHisHis---ProProGlnHisProProThrPro 110  
 QY 494 AATCATATGGTCCAGCTAACAGATC-TGAGGGCAACTGCTGTCTAGTGGCCAGGGCTG 552  
 Db 111 ThrThrThrProProHisHisAla-----ProThrPro 122  
 QY 553 CACCTGCCATCCCGGCTCTGCCACTTTAGGCGCTCTTAGAGCAGTGTCTTAGGAAGT 612  
 Db 123 HisThrProProThrThrPro----- 130

QY 613 ACCTGAGGATGEGTTTTCTGTCTCTGTGTACAGGGCAGCTGATGGGATAAGGTGGGAA 672  
 Db 130 -----ProArgProProThrThrHisThrHisThrProProHisPro 130  
 QY 673 GGACGGTCACTGCTTGGGCCAGCTGCCAGCCCTGGGATGGGAAACCAACCATGTC 732  
 Db 131 -----ProArgProProThrThrHisThrHisThrHisThrHisThrHisThr 144  
 QY 733 CCCAGCAGAGGGCCAGAGTGGAACTGCTCAT-----GCCCTTCGT 777  
 Db 145 ProThrProProProLeuProThrThrProHisProHisProHisSerHisSerThrLeuSer 164  
 QY 778 CTTGAGGAGCTGAGTGGGCGACAGGGCCAGGGGAAGTTTTTCAGGCCTTCATCAAAG 837  
 Db 165 ProHisHisPro-----His 169  
 QY 838 AGAACACATCTCAGCTCCGACCCCTCATCTCTGTATCAGCACTTACCGGTGTGACT 897  
 Db 170 SerThrThrSerSerLeuPro-----SerThrHisAsnAsnIleThr 183  
 QY 898 GCCCTGTCTAGCTAGCATACGGTGGGCCCA 927  
 Db 184 AsnThrProProAlaHisThrLeuThrPro 193

RESULT 22  
 US-09-989-920-199  
 ; Sequence 199, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Hervé  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 199  
 ; LENGTH: 162  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-199

Alignment Scores:  
 Pred. No.: 8.9 Length: 162  
 Score: 56.50 Matches: 38  
 Percent Similarity: 40.46% Conservative: 15  
 Best Local Similarity: 29.01% Mismatches: 37  
 Query Match: 1.11% Indels: 42  
 DB: 1 Gaps: 7

us-09-989-920-100 (1-2754) x US-09-989-920-199 (1-162)  
 QY 1377 CTTGTGCTGGGAGCCCTCCAGGTTCTGGTCTTCCAGTGGCTCTCGCAGAG 1436  
 Db 17 ProHisThrProAlaSer----- 22  
 QY 1437 TCCACCCGCTGCTGGTGGAAATGCAGACCTTTCTTCTTCTTCCCTCCCTCCCTC 1496  
 Db 23 SerAlaAlaProHis-LeuSerLeuPheSerProLysLeuValPheLeu-ThrIleIleV 42  
 QY 1497 CTGTTCTGGGACCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1556  
 Db 42 alValGlyGly-----GlyGlnMetLeuLys----- 50  
 QY 1557 GATCGTGGAGAGGACTTTGAGCAGTGGTGGCAGCAGTGGCC-----TCCT 1604  
 Db 51 -----ValGluAlaAspLeuGluLysGluThrHisGlyValThrValAlaLysAspSer 69





TYPE: PRT
ORGANISM: Homo sapien
US-09-989-920-200

Alignment Scores: 11.7 Length: 594
Pred. No.: 55.00 Matches: 47
Score: 35.53% Conservative: 23
Percent Similarity: 23.86% Mismatches: 69
Best Local Similarity: 23.86% Indels: 59
Query Match: 1.08% Gaps: 10
DB: 10

us-09-989-920-100 (1-2754) x US-09-989-920-200 (1-594)

1433 AGAGTCACCCCGCTCGTGGTGGAAATGCAGAGCCCTTGTCTTCTTCCCTTCCGCGCTG 1492
268 ArgPheHisProGluLeuTrp-----CysSerGly 277
1493 CTTCCCTTCTGGGACCCCGCTGGGCTTGTCTGCATCCCTCCGCGAGTCCCTCAG 1552
278 ArgSerValProLeuAspArgGlnGlyTyr-----GlyGlnIleLysVal 292
1553 GG-TTGATGCGTGAGAAAGACTTTGACAGTGTGGGCGACAGTGGCCCTC----- 1602
293 ValArgAlaAspGlyAspThrLeuSerCysIleCysGlyLysThrLysLeuGlyGluAsp 312
1603 ---CTGGCCAGCTCACACTCTTGTCTGGAGGGGCGAGCCTGATCTCACCTCCACCTAGT 1659
313 MetLeuCysLeuLeuHis-----GlyArgAsnSerMet---AlaProProCysGly 328
1660 ACCTTGGGAGTGGACCTTTGGCTTCTGTGGAGCCTCAAGCCTTCCCATGTGTC 1719
329 AspMet-----GluAsnLeuLeuCysAlaThrAsp-----SerLeuTyr----- 341
1720 CAGCTGTCTTCTGCTACAAGGGGACTGCTCACAGTGGCCCTCAGCTGGTGGTTTGA 1779
342 -----LeuAspThrMetGlnValMetLysTrpPheGln 352
1780 GGGCGCGCCCGCCCTCCATAAGGATATCTTGGGCGCTGAGAAATCTGCATCTGCCAT 1839
353 ThrAlaLeuThrArgAlaTrpLysGlyIleAlaHisLysTyrGluPheAspLeuAlaPhe 372
1840 TGGAGGATGGAC-----AGCCTCAAATGGAGGAGTCCACGGAGATGGTCCGA 1890
373 GlyGlnLeuAspSerProGlySerLeuLysIleLysPheArgSerGlyLysPheMetPro 392
1891 GGTCCGGCTGTGGCCATC----- 1908
393 PheAsnLeuIleProValIleGlnCysAspAspSerAspLeuTyrPheValSerHisLeu 412
1909 -----CAGCCCGCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGTGTCTT 1953
413 ProArgGluProSerGluGlyThrProAlaSerSerThrAspTrpLeuLeu 429

RESULT 25
US-09-989-920-249
Sequence 249, Application US/09989920
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Herve
APPLICANT: Chen, Sei-Yu
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
FILE REFERENCE: DEX-0291
CURRENT APPLICATION NUMBER: US/09/989, 920
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/252,500
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 249
LENGTH: 594

TYPE: PRT
ORGANISM: Homo sapien
US-09-989-920-249

Alignment Scores: 11.7 Length: 594
Pred. No.: 55.00 Matches: 47
Score: 35.53% Conservative: 23
Percent Similarity: 23.86% Mismatches: 69
Best Local Similarity: 23.86% Indels: 59
Query Match: 1.08% Gaps: 10
DB: 10

us-09-989-920-100 (1-2754) x US-09-989-920-249 (1-594)

1433 AGAGTCACCCCGCTCGTGGTGGAAATGCAGAGCCCTTGTCTTCTTCCCTTCCGCGCTG 1492
268 ArgPheHisProGluLeuTrp-----CysSerGly 277
1493 CTTCCCTTCTGGGACCCCGCTGGGCTTGTCTGCATCCCTCCGCGAGTCCCTCAG 1552
278 ArgSerValProLeuAspArgGlnGlyTyr-----GlyGlnIleLysVal 292
1553 GG-TTGATGCGTGAGAAAGACTTTGACAGTGTGGGCGACAGTGGCCCTC----- 1602
293 ValArgAlaAspGlyAspThrLeuSerCysIleCysGlyLysThrLysLeuGlyGluAsp 312
1603 ---CTGGCCAGCTCACACTCTTGTCTGGAGGGGCGAGCCTGATCTCACCTCCACCTAGT 1659
313 MetLeuCysLeuLeuHis-----GlyArgAsnSerMet---AlaProProCysGly 328
1660 ACCTTGGGAGTGGACCTTTGGCTTCTGTGGAGCCTCAAGCCTTCCCATGTGTC 1719
329 AspMet-----GluAsnLeuLeuCysAlaThrAsp-----SerLeuTyr----- 341
1720 CAGCTGTCTTCTGCTACAAGGGGACTGCTCACAGTGGCCCTCAGCTGGTGGTTTGA 1779
342 -----LeuAspThrMetGlnValMetLysTrpPheGln 352
1780 GGGCGCGCCCGCCCTCCATAAGGATATCTTGGGCGCTGAGAAATCTGCATCTGCCAT 1839
353 ThrAlaLeuThrArgAlaTrpLysGlyIleAlaHisLysTyrGluPheAspLeuAlaPhe 372
1840 TGGAGGATGGAC-----AGCCTCAAATGGAGGAGTCCACGGAGATGGTCCGA 1890
373 GlyGlnLeuAspSerProGlySerLeuLysIleLysPheArgSerGlyLysPheMetPro 392
1891 GGTCCGGCTGTGGCCATC----- 1908
393 PheAsnLeuIleProValIleGlnCysAspAspSerAspLeuTyrPheValSerHisLeu 412
1909 -----CAGCCCGCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGTGTCTT 1953
413 ProArgGluProSerGluGlyThrProAlaSerSerThrAspTrpLeuLeu 429

RESULT 27
US-09-989-920-274
Sequence 274, Application US/09989920
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Herve
APPLICANT: Chen, Sei-Yu
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
FILE REFERENCE: DEX-0291
CURRENT APPLICATION NUMBER: US/09/989, 920
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/252,500
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 274
LENGTH: 224

TYPE: PRT
ORGANISM: Homo sapien
US-09-989-920-274

Alignment Scores:
Pred. No.: 12.4
Score: 54.50
Percent Similarity: 33.96%
Best Local Similarity: 25.16%
Query Match: 1.10%
DB: 1

US-09-989-920-100 (1-2754) x US-09-989-920-274 (1-224)
QY 1960 TGGAGTGAAGACACACAGGGGTCACAGAGGCTGGCAAGCCACAGGGGCTGGATGGCCA 1901
Db 94 TTPSerLeuAspValAsnGly-----GlyAspGluTrpLys 105
QY 1900 CAGCCGGACCTGGACCCATCTCCGTTGGGACTCTTCCATTTGAGGCTGCCATCTCC 1841
Db 106 ValGluAspLeuSerArgAspGlnArgLysGlu-----PheProAsnAsp 120
QY 1840 AATGGCAGATGAGATT---CTCAGGCCAGGATACCCCTTATGAGGGCCCGGGGGCGG 1784
Db 121 GlnValArgSerGlnAlaArgLeuArgValGlnValProAlaValArg----- 136
QY 1783 CCCTCAAACACACAGCTGAGCCACTGTGAGCAGTCCCTTTGTAGCAGGAAGACA 1724
Db 137 -----SerAlaProValValArgAlaArgAla 145
QY 1723 GCTGGACACATGGAGAGGCTTGCAGGCTCCAGAGAAGCCA-----AAGGTCT 1673
Db 146 SerGlyAspLeu-----ProAlaArgProGlyAspHisProAlaGlu 159
QY 1672 CAGTCCCC-AAGTACTAGTGGAGTGGATCAGCAGCTCCCTCCAGGACAGAGTGT 1614
Db 160 GluArgCysGlnValGluGlyLeuProHisIleLeuGlnLeuPro----- 175
QY 1613 GAGCTGCCAGGAGGCCACTGTGCCACCCTGCTCAAAGTCTTCTCCAGC----- 1560
Db 176 -----AlaArgArgProLeuHisLeuValSerAlaArgArgGlyHisSerLeuLeu 193
QY 1559 -----ATCAACCTGAGGACCTGGCCAGGGGATGCAGACCAAGGCCCGC 1512
Db 194 GlyArgLeuValArgProGluGlyHisGlnGlnGlnHisHisIleSargAlaProAla 212

RESULT 28
US-09-989-920-264
Sequence 264, Application US/09989920
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Hervé
APPLICANT: Chen, Sei-Yu
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
FILE REFERENCE: DEX-0291
CURRENT APPLICATION NUMBER: US/09/989,920
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/252,500
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 264
LENGTH: 100
MATCHES: 54.00
TYPE: PRT
ORGANISM: Homo sapien
US-09-989-920-264
Alignment Scores:
Pred. No.: 12.7
Score: 54.00
Percent Similarity: 41.03%

Best Local Similarity: 30.77%
Query Match: 1.06%
DB: 1
us-09-989-920-100 (1-2754) x US-09-989-920-264 (1-100)
QY 1679 TTTTGGCTTCTGGAGCCTGCAAGCCCTTCCATGTCCAGCTGTCCAGTCTCTTCCCTGCTAC 1738
Db 1 PhePheLeuArgTrpSerLeuAlaGlnValAlaGlnAlaAla----- 14
QY 1739 AAAGGGACTGCTCACAGTGGCCTCAGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1795
Db 15 -----ArgGlnTrpLeuAsnLeuSerSerLeuGlnProProProProGly 29
RESULT 29
US-09-989-920-247
Sequence 247, Application US/09989920
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Hervé
APPLICANT: Chen, Sei-Yu
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
FILE REFERENCE: DEX-0291
CURRENT APPLICATION NUMBER: US/09/989,920
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/252,500
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 247
LENGTH: 240
MATCHES: 16.9
TYPE: PRT
ORGANISM: Homo sapien
US-09-989-920-247
Alignment Scores:
Pred. No.: 16.9
Score: 52.50
Percent Similarity: 35.68%
Best Local Similarity: 23.78%
Query Match: 1.03%
DB: 1
us-09-989-920-100 (1-2754) x US-09-989-920-247 (1-240)
QY 1932 CTCTGTGCACCCCTGGTGTCTTCTCCAGGGGACAGCAGCCACTGCAGTCTCTTCTTC 1991
Db 21 ValCysAlaThrLeuIleLeuLeuLeuGlyLeuSerGlyLeuGlyLeuGlySerPhe 40
QY 1932 TTCGTGAGTAC-----AGTAGTGCAT----- 2012
Db 41 LeuLeuThrHisArgThrGlyLeuArgSerProAspIleProGlnAspTrpValSerPhe 60
QY 2013 ---AGCAGCTGGGGCTAACAGGCTAGGCTTTGT-----GTTCTCGGCATT 2054
Db 61 LeuArgSerPheGly---GlnLeuThrLeuCysProArgAsnGlyThrValThrGlyLys 79
QY 2055 TGGTCAGCTTCTCCTCGATC-----CTCCCTAAAGCAATGGG 2093
Db 80 TrpArgGlySerHisValValGlyLeuLeuThrThrLeuAsnPheGlyAspGlyProAsp 99
QY 2094 AGGCCCCACTAGCCCAAGTTCCTCAGGAAGTCACTGGAGGTTAGATGGGGCCAGG--- 2150
Db 100 ArgAsnLysThrArgThrPheGlnAlaThrValLeuGlySerGlnMetGlyLeuLysGly 119
QY 2151 -----GTCCACAGCTACTGATGGCCCGAGCCAGGTTGAG----- 2185
Db 120 SerSerAlaGlyGlnLeuValIleThrAlaArgValThrGluArgThrGluArgThrAlaGly 139
QY 2186 -----CTTCTGGTGTCCAGTCCCGGATCCACTCCAGATCTCA 2224
Db -----

Alignment Scores:
Pred. No.: 12.7
Score: 54.00
Percent Similarity: 41.03%

Db 140 ThrCysLeuTyPheSerAlaValProGlyLeuProSerSerGlnProIleSer 159  
 QY 2225 TGCTCTCAGATAGGGGACAGTCTTTTTCACAGGCTGGCTCTCTCTGAGGCTC 2284  
 Db 160 CysSerGluGluGlyAlaGlyAsnAlaThrLeuSerProArg----- 173  
 QY 2285 ATTGCTGGCTGGCTCTCTCTGGAAAGCTTTGGGGGCTTGTCT-----TGG 2335  
 Db 174 ---MetGlyGluGluCysValSerValTrpSerHisGluGlyLeuValLeuThrLysLeu 192  
 QY 2336 TTAACACACAGAG 2350  
 Db 193 LeuThrSerGluGlu 197

RESULT 30  
 US-09-989-920-208  
 ; Sequence 208, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; PRIORITY FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 208  
 ; LENGTH: 130  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-208

Alignment Scores:  
 Pred. No.: 17.7 Length: 130  
 Score: 52.00 Matches: 14  
 Percent Similarity: 43.18% Conservative: 5  
 Best Local Similarity: 31.82% Mismatches: 17  
 Query Match: 1.05% Indels: 8  
 Gaps: 1  
 us-09-989-920-100 (1-2754) x US-09-989-920-208 (1-130)

QY 725 TTGGTTTCCCATCCGACGGCTGGCCAGCTGGGCCCAACGACTGACCGTCTTCCCA 666  
 Db 53 PheGlyPheProVal-----GluLeuProArgProGlyPro 64  
 QY 665 CCTATCCCATCAGTCCCTGCACAGGACGAGAAACCCTATGCTCAGACTACTTCT 606  
 Db 65 ThrGlyAlaTyLysLysValLysAsnGlnAsnGlnThrThrSerSerGluLeuLeuArg 84  
 QY 605 AAGGACACTGCC 594  
 Db 85 LysGlnThrSer 88

RESULT 31  
 US-09-989-920-185  
 ; Sequence 185, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500

; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 185  
 ; LENGTH: 489  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-185

Alignment Scores:  
 Pred. No.: 17.8 Length: 489  
 Score: 52.00 Matches: 17  
 Percent Similarity: 44.44% Conservative: 7  
 Best Local Similarity: 31.48% Mismatches: 22  
 Query Match: 1.02% Indels: 9  
 Gaps: 2  
 us-09-989-920-100 (1-2754) x US-09-989-920-185 (1-489)

QY 2144 GGCCAGGTCACACAGTCTACTGATGGCCGAGCAGGTTGAGCTTCCGTTCCAGTCC 2203  
 Db 27 GlyAsnValAlaThrAsnThrAspGlyAsnTyrCysGlyLeuProGlyAsnGluGln 46  
 QY 2204 GGATCCCACTTCAGAT-CTCATGCTCTCAGATAGTGGGACAAAGTTCTTTTTCACAGT 2262  
 Db 47 AlaCysLysIleLysSerPheTyLeu-----LysTrpAspPhePhe----- 60  
 QY 2263 GCTGGCTCTGCTCCTCAGGCTCCTCATTTGGCTGGTGGTGGTCTC 2304  
 Db 61 -----AlaLeuLysAsnIleHisCysTrpLysProValLeu 72

RESULT 32  
 US-09-989-920-212  
 ; Sequence 212, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; PRIOR FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 212  
 ; LENGTH: 610  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-212

Alignment Scores:  
 Pred. No.: 18.6 Length: 610  
 Score: 51.50 Matches: 49  
 Percent Similarity: 30.16% Conservative: 27  
 Best Local Similarity: 19.44% Mismatches: 83  
 Query Match: 1.04% Indels: 93  
 Gaps: 11  
 us-09-989-920-100 (1-2754) x US-09-989-920-212 (1-610)

QY 1962 CTGGAGTGAACACCCAGGGGTG-----CACAGAGCTGGACCAAGCCACAGG 1915  
 Db 382 ProGly---LysThrProHisThrCysAspGluCysGlyLysAlaPhePheSerSerArg 400  
 QY 1914 GGCTGGATGCCACAGCCGACCTCGACCCATCTCCCGTGGACTCCTTCCATTTGAG 1955  
 Db 401 ThrLeuIleSerHisLysArgValHis-----LeuGlyGluLysProPheLys 416  
 QY 1854 GCTGTTCATCTCCATGGCAGATGTCAGAAATTCAGGCCCCAGGATACCCCTTATGAGGG 1795

Db 417 CysValGlu-----CysGlySerPheSerTyr--- 426  
 Qy 1794 CCGGGGGGGCCCTCAAACCCAGCTGAGGACCTGAGAGTCCCTTTGTAG 1735  
 Db 427 -----SerSerLeuLeuSer 431  
 Qy 1734 CAGGAAGAGCAGCTGGACATGGGAAGAGGCTTGCAGGCTCCAGAGAAGCCAAAGAGTC 1675  
 Db 432 GlnHisysArgIleHisThrGlyGluYbProTyrValCysAspArgCys----- 448  
 Qy 1674 CTCAGTCCCCAGGTACTAGTGGAGTGCAGTGCAGCTGCCCTCCAGACAGAGTG 1615  
 Db 449 -----GlyLysAlaPheArgAsnSerSerGlyLeuThrValHisLysArgIle 464  
 Qy 1614 TGAGCTGGCAGGAGGACCTGCTGCCACCACTGCTCAAGTCTTCCACGATCAA 1555  
 Db 465 HisThrGlyGluYbProTyrGluCysAspGluCysGlyLysAlaTyrIleSerHis--- 483  
 Qy 1554 CCCTGAGGACCTGGCCAGGGATGCAGACCAAGCCCGGGTCCCGAGAACAGGA 1495  
 Db 484 -----SerSerLeuIleAsnHis 489  
 Qy 1494 AGCAGGGCCGAAGAAAGCAA--- 1471  
 Db 490 LysSerValHisGlnGlyLysGlnProTyrAsnCysGluCysGlyLysSerPheAsnTyr 509  
 Qy 1470 AGGCTGTGATCCACCACGAGGGGGTGGCTGCAGGAGCCAGCTGAGCACCA 1411  
 Db 510 ArgSerValLeuAspGlnHisLysArgIleHisThrGlyLysPro----- 525  
 Qy 1410 GAACCTTCCAAAGGGGCTCCCCAGCCACACAGGACAGGATGCTGATAGCCAGGGTT 1351  
 Db 526 -----TyrArgCys----- 528  
 Qy 1350 AATCAGAAACCGCTCATCTGATGCTCCCTCCAGGGCAACCAAGTATTT 1291  
 Db 529 ---AsnGluCysAlaHisIleProAsnAlaThrAlaAspLeuMetLysValAspHis--- 546  
 Qy 1290 ATTCTTTTAAAGCCCTTACTGCTGAGGCTTC 1255  
 Db 547 -----GluGluGluProGlnLeuSerGluProTyr 556

RESULT 33  
 US-09-989-920-230  
 ; Sequence 230, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 230  
 ; LENGTH: 53  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-230  
 Alignment Scores:  
 Pred. No.: 20.9 Length: 53  
 Score: 50.50 Matches: 15  
 Percent Similarity: 53.85% Conservative: 6  
 Best Local Similarity: 38.46% Mismatches: 11  
 Query Match: 1.02% Indels: 7  
 Gaps: 1

us-09-989-920-100 (1-2754) x US-09-989-920-230 (1-53)  
 Qy 721 GTTTCCTCCATFCGCCAGCTGGCCAGCTGGGGCCCAAGCACTGACCGTCTCTCCACCTT 662  
 Db 14 LeuGlnProAlaProSerTrpLeu-----LysGlnAlaLeu-----HisLeu 27  
 Qy 661 ATCCCA---TCAGTCCCTGCACAGGAGCAGAAAAACCATCCCTCAGAGCTACTTC 608  
 Db 28 SerProLeuSerSerAlaHisTyrArgHisThrProHisProAlaAsnPhePhe 46  
 RESULT 34  
 US-09-989-920-184  
 ; Sequence 184, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: 60/252,500  
 ; NUMBER OF SEQ ID NOS: 284  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 184  
 ; LENGTH: 101  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-989-920-184  
 Alignment Scores:  
 Pred. No.: 27.6 Length: 101  
 Score: 49.00 Matches: 25  
 Percent Similarity: 40.51% Conservative: 7  
 Best Local Similarity: 31.65% Mismatches: 29  
 Query Match: 0.99% Indels: 19  
 Gaps: 5

us-09-989-920-100 (1-2754) x US-09-989-920-184 (1-101)  
 Qy 2673 CAGACTTCTGTCGG-----CACTATGCTACAGTAAACAGTGCACAAATATGCTTTA 2623  
 Db 7 GlnAlaProArgProPheLeuTyrHisGlyCysTrpValThrSerGlySerHisHisLeu 26  
 Qy 2622 TTT-----CAGGTACAAAAACATGGCAGTAG-GACAACCTTTGAG 2585  
 Db 27 PheProSerLeuPheProIleSerGlnMetTrpGlyHisGlyLeuAspGlyLeuHis 46  
 Qy 2584 CTCAGCCCACTC-----CCAGGACTCTCACCCACCTGACCCAT 2540  
 Db 47 ArgSerPheHisLeuCysGluSerLysSerGlyGlnSerAlaArgHisLeuCys--- 65  
 Qy 2539 CCAAGGGCCACACCCCGCAGATACTCCACCCCTTTAGAAAAGAGTCAACCA 2483  
 Db 66 ProGlySerAlaProGlnAsnGln-----ProProAlaSerLeuLysGlnLysPro 82  
 RESULT 35  
 US-09-989-920-220  
 ; Sequence 220, Application US/09989920  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-Yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot  
 ; FILE REFERENCE: DEX-0291  
 ; CURRENT APPLICATION NUMBER: US/09/989,920  
 ; CURRENT FILING DATE: 2001-11-21





```

Query Match: 0.94% Indels: 28
DB: 1 Gaps: 4
us-09-989-920-100 (1-2754) x US-09-989-920-259 (1-1533)
QY 1423 TGGCTCCAGAGTCCACCCCGCTCGTGTGGAAATGCAGAGCCCTTCTTCCTTC 1482
Db 22 TrpGlnLeuGluLeu---ProLysLeuLeuIleSerValHisGlyLeuGlnAsnPhe 40
QY 1483 -----TTGCCCCCTCTTCCTGTTCC 1503
Db 41 GluMetGlnProLysLeuLysGlnValPheGlyLysGlyLeuIleLysAlaAlaMetThr 60
QY 1504 TGGGACCCGCTGGGCCCTTGTCTGTCATCCCTCCAGGTCCTCAGGTTGATGCGT 1563
Db 61 ThrGly-----AlaTrpIlePheThrGlyValSerThr 72
QY 1564 GGAGAAGGACTTTGAGCAGTGTGGCAGCAGTGGCCTCCCTGGCAGCTCACACTTGT 1623
Db 73 Gly-----ValIleSerHisValGlyAspAlaLeuLysAspHisSerSer 87
QY 1624 CCTGGAGGGGAGCTGATCTCACCTCCACTAGTACTCTGGGACTG 1672
Db 88 LysSerArgGlyArgValCysAla-IleGlyIleAlaProTrpGlyIle 103
RESULT 41
US-09-989-920-279
; Sequence 279, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto Length: 149
; APPLICANT: Recipon, Herve Matches: 24
; APPLICANT: Chen, Sei-Yu Conservative: 5
; APPLICANT: Sun, Yongming Mismatches: 23
; APPLICANT: Liu, Chenghua Indels: 36
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins FILE REFERENCE: DEX-0291
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 279
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-279
Alignment Scores:
Pred. No.: 37.6 Length: 149
Score: 47.00 Matches: 24
Percent Similarity: 32.95% Conservative: 5
Best Local Similarity: 27.27% Mismatches: 23
Query Match: 0.93% Indels: 36
DB: 1 Gaps: 6
us-09-989-920-100 (1-2754) x US-09-989-920-279 (1-149)
QY 1336 AGCGTTTCATTAACCCCTGCTATCAGCATCTGCTGCTGCTGCTGCTGCTGCTG 1395
Db 26 SerArgValCysLeuArgLeuValLeuSerTrpSerArgValValCysPheTrp----- 43
QY 1396 CCCCTTGGAGCTTCTGCTGCTTCCAGTGGCTCCCTGAGATCCACCCCGCTGGGTG 1455
Db 44 -----TrpSerPhe---TrpLeuPheValSer-----ValValCys 54
QY 1456 GGAATGCAGAGCCCTTTCCTTCTTCTTGGCCGCTGCTCTCTCTCTCTCTCTCT 1500
Db 55 PheValPheSerCysPheValSerLeuLeu-----CysCysCysGlyValArgLeuTyr 72
QY 1501 -----TCTGGGGACCC----- 1512
Db 73 PheValValSerTrpGlyValPhePheCysAspLeuLeuArgCysCysTyrAspAsnVal 92

```

```

QY 1513 -----GCTGGGCTTGGTCTGC 1530
Db 93 CysPheAlaHisProThrValCys 100
RESULT 42
US-09-989-920-187
; Sequence 187, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins FILE REFERENCE: DEX-0291
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 187
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-187
Alignment Scores:
Pred. No.: 40.4 Length: 149
Score: 46.50 Matches: 28
Percent Similarity: 37.86% Conservative: 11
Best Local Similarity: 27.18% Mismatches: 27
Query Match: 0.92% Indels: 37
DB: 1 Gaps: 5
us-09-989-920-100 (1-2754) x US-09-989-920-187 (1-149)
QY 705 CTTGGATGGGAAACCAACCACTGTCCTCCAGAGGCGCCAGAGTGGAACTGTC 764
Db 1 ProGlyAsnLeuAspThrSer-----SerArgGlySerSerGlySerProAla 16
QY 765 TCATGCCCTTCCTCCAGGAGCCCTGAGTGGCCAGCAG----- 804
Db 17 -HisAlaGluSerTyrSerSerGlyGlyGlyGlnGlnLysPheArgValAspMetPr 36
QY 805 -GGCCAGGGAAGTTTTCAGGCTTCATCAAGAGAACCAACCACTGTCCTCCAGCT 855
Db 36 oGlySerGlySerAlaPheIleProThrIleAsnAlaIleThrThrSerGlnAspLeuG1 56
QY 856 -----CCGCACCCCTCATC 869
Db 56 nTrpMetValGlnProThrValIleThrSerMetSerAsnProTyrProArg-SerHisP 76
QY 870 CTGTATCAGCACTTACCGTGTGTGACTGCTCCCTGTCAGCTAGCATACGGTGGGCCACC 929
Db 76 roTyrSerProLeuProGlyLeu-----AlaSerValAla---GlyHisM 90
QY 930 TGGCC 934
Db 90 etAla 91
RESULT 43
US-09-989-920-233
; Sequence 233, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins FILE REFERENCE: DEX-0291

```







Ac079988 (qi 10/13/20)





Exhibit B (page 1 of 2)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 08:42:07 ; Search time 227 Seconds  
(without alignments)  
8623.403 Million cell updates/sec

Title: US-09-989-920-100  
Perfect score: 2754  
Sequence: 1 gccagagagcctcagctt.....aaaaaaaaaagatgggcc 2754

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database :
- 1: /cgm2\_6/ptodata/1/ina/5A\_COMB.seq.\*
  - 2: /cgm2\_6/ptodata/1/ina/5B\_COMB.seq.\*
  - 3: /cgm2\_6/ptodata/1/ina/6A\_COMB.seq.\*
  - 4: /cgm2\_6/ptodata/1/ina/6B\_COMB.seq.\*
  - 5: /cgm2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*
  - 6: /cgm2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1747.6	63.5	1853	3	US-09-232-160-3
2	52	1.9	7218	1	US-08-232-463-14
3	49	1.8	6330	3	US-09-880-427-2
4	49	1.8	6330	3	US-09-306-538B-2
5	49	1.8	18806	4	US-09-748-028A-3
6	49	1.8	19806	4	US-10-118-037-3
7	47	1.7	265	4	US-09-513-999C-27278
8	42.6	1.5	640681	4	US-09-790-988-1
9	41.6	1.5	210	4	US-09-248-796A-12137
10	41.4	1.5	329	4	US-09-513-999C-10733
11	40.8	1.5	318	4	US-09-513-999C-11322
12	40.8	1.5	398	4	US-09-513-999C-23077
13	40.4	1.5	1141	4	US-09-808-708B-22
14	40.4	1.5	640681	4	US-09-790-988-1
15	39.6	1.4	1107	4	US-09-248-796A-8906
16	39.4	1.4	588	3	US-09-328-111-321
17	39.4	1.4	5852	3	US-07-867-106-2
18	39	1.4	303	4	US-09-601-198-179
19	38.8	1.4	9636	1	US-08-323-170B-1
20	38.8	1.4	9636	3	US-08-954-441-1
21	38.4	1.4	2422	1	US-07-867-106-5
22	38.4	1.4	5852	1	US-07-867-106-2
23	38.4	1.4	54945	4	US-09-967-669-10
24	38.2	1.4	1099	4	US-09-270-767-11371
25	38.2	1.4	3009	4	US-09-248-796A-6806
26	38.2	1.4	3454	4	US-09-270-767-11909
27	38.2	1.4	5185	3	US-08-971-393-5

c 28	38.2	1.4	8561	3	US-09-112-450-3	Sequence 3, Appli
c 29	38.2	1.4	8561	4	US-09-419-291A-3	Sequence 3, Appli
c 30	38.2	1.4	98844	4	US-09-791-211-10	Sequence 10, Appli
c 31	38	1.4	2007	2	US-08-743-637B-169	Sequence 169, App
c 32	38	1.4	2007	3	US-08-526-840B-169	Sequence 169, App
c 33	38	1.4	2028	3	US-09-134-001C-1710	Sequence 2, Appli
c 34	37.8	1.4	4140	3	US-08-894-731-2	Sequence 1, Appli
c 35	37.8	1.4	8920	2	US-08-446-855A-1	Sequence 1, Appli
c 36	37.8	1.4	8920	3	US-08-150-741-1	Sequence 1, Appli
c 37	37.8	1.4	1664976	4	US-08-316-421B-1	Sequence 1, Appli
c 38	37.8	1.4	1664976	4	US-09-692-570-1	Sequence 3, Appli
c 39	37.4	1.4	2663	1	US-08-136-743B-3	Sequence 70, Appli
c 40	37.4	1.4	6040	4	US-10-204-708-70	Sequence 3, Appli
c 41	37.4	1.4	74962	4	US-09-685-853A-3	Sequence 3767, Ap
c 42	37.2	1.4	942	4	US-09-270-767-3767	Sequence 19049, A
c 43	37.2	1.4	942	4	US-09-270-767-19049	Sequence 5, Appli
c 44	37.2	1.4	1928	4	US-09-674-826B-5	Sequence 4, Appli
c 45	36.8	1.3	80246	3	US-09-078-294-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
 US-09-232-160-3  
 ; Sequence 3, Application US/09232160  
 ; Patent No. 6368794  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Steve Daniel  
 ; APPLICANT: James Gilmore  
 ; APPLICANT: Susan G. Stuart  
 ; APPLICANT: Laura Stuve  
 ; TITLE OF INVENTION: DETECTION OF ALTERED EXPRESSION OF GENES REGULATING CELL  
 ; FILE OF INVENTION: PROLIFERATION  
 ; FILE REFERENCE: PA-0003 US  
 ; CURRENT APPLICATION NUMBER: US/09/232,160  
 ; CURRENT FILING DATE: 1999-01-15  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 3  
 ; LENGTH: 1853  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: 1283330  
 ; US-09-232-160-3

Query Match	Score	Length	DB 3;	Length	1853;
Best Local Similarity	63.5%	1747.6;	DB 3;	Length	1853;
Matches 1750;	Conservative	0;	Mismatches	4;	Indels 0;
Gaps	0;				
Qy	994	TGTTTTTCTCCCCATCACAAAAAATCTTATTTTTAGTAGACATGATTTTACCAA	1053		
Db	1	TGTTTTTCTCCCCATCACAAAAAATCTTATTTTTAGTAGACATGATTTTACCAA	60		
Qy	1054	AAATGTACTCAATTTGATTTTTGGATTTTATCAATTTAAAAAATGTGGAAATTTG	1113		
Db	61	AAATGTACTCAATTTGATTTTTGGATTTTATCAATTTAAAAAATGTGGAAATTTG	120		
Qy	1114	TTGCTTTACGCCACATTAATTCATTTTCCCTTGGCTTGAAGCCCAATATTT	1173		
Db	121	TTGCTTTACGCCACATTAATTCATTTTCCCTTGGCTTGAAGCCCAATATTT	180		
Qy	1174	ACCGTCTAGCCCTTACAGAAAAGTCTGCTACTGAGCCAGACCTTACCTCC	1233		
Db	181	ACCGTCTAGCCCTTACAGAAAAGTCTGCTACTGAGCCAGACCTTACCTCC	240		
Qy	1234	ATCCCTGTGGATTTTAAAGAAAGCCTCAGACAGTAGGGCTTTTAAAGATAAA	1293		
Db	241	ATCCCTGTGGATTTTAAAGAAAGCCTCAGACAGTAGGGCTTTTAAAGATAAA	300		
Qy	1294	ATGACTTGGTTTGGCTTGAAGCCAGGGAACCTTACAGTGGGGTTTCTGCAATTAAC	1353		

301 ATGACTTGGTTGGTGGAGACGGGGAAGCAITTCAGATGAGCGGTTTCGTCAATTAAC 360  
 1354 CTTGSCATATACGCACTCTGTFCTGTTGCTGGCGAGCCCTTGGAGGTTCTGG 1413  
 361 CTTGSCATATACGCACTCTGTFCTGTTGCTGGCGAGCCCTTGGAGGTTCTGG 420  
 1414 TGCCTCAGTGGTCTGTCAGAGTCCACCCGCTGCTGGTGGGATGAGAGCCCTTTG 1473  
 421 TGCCTCAGTGGTCTGTCAGAGTCCACCCGCTGCTGGTGGGATGAGAGCCCTTTG 480  
 1474 CTTTCCCTTCTGCGGCTGTTCTCTGTTCTGTTGGGACCCGCTGGGCTTTGTTGTCATC 1533  
 481 CTTTCCCTTCTGCGGCTGTTCTCTGTTCTGTTGGGACCCGCTGGGCTTTGTTGTCATC 540  
 1534 CCTTGGCAGGTCCTCAGAGTTCAGTCCGTTGGAGAGGACTTTGACGATGTTGGGCGAG 1593  
 541 CCTTGGCAGGTCCTCAGAGTTCAGTCCGTTGGAGAGGACTTTGACGATGTTGGGCGAG 600  
 1594 AGTGGCTCCTGGCCAGCTCACACTCTTCTGTTGGGAGGGGAGCCCTGATCTCACTCCA 1653  
 601 AGTGGCTCCTGGCCAGCTCACACTCTTCTGTTGGGAGGGGAGCCCTGATCTCACTCCA 660  
 1654 CTTAGTACCTTGGGACTGAGGACTTTGGCTTCTGTTGGGAGGGGAGCCCTGATCTCCA 1713  
 661 CTTAGTACCTTGGGACTGAGGACTTTGGCTTCTGTTGGGAGGGGAGCCCTTCCA 720  
 1714 TGTGTCAGTCTCTTCTGTTACAAAGGGGACTGCTCACAGTGGCCCTCAGCTTGGTGG 1773  
 721 TGTGTCAGTCTCTTCTGTTACAAAGGGGACTGCTCACAGTGGCCCTCAGCTTGGTGG 780  
 1774 TTTGAGGGCCGCCCCCGCCCTCATAAGGGATCTCTGGCTGAGAAATTCGATCTC 1833  
 781 TTTGAGGGCCGCCCCCGCCCTCATAAGGGATCTCTGGCTGAGAAATTCGATCTC 840  
 1834 TGCATGAGGATGAGCAGCCCTCAATGGAAGGAGTCCACGGGAGTGGTCCGAGGT 1893  
 841 TGCATGAGGATGAGCAGCCCTCAATGGAAGGAGTCCACGGGAGTGGTCCGAGGT 900  
 1894 CCGCTGTGGCATTCAGACCCCTGTTGGCTTCTGTTGGGAGGGGAGTGGTCCGAGGT 1953  
 901 CCGCTGTGGCATTCAGACCCCTGTTGGCTTCTGTTGGGAGGGGAGTGGTCCGAGGT 960  
 1954 CACTCCAGGGGACAGCAGCAGCCCTGAGTCTCTTCTGTTGGGAGTGGTCCGAGGT 2013  
 961 CACTCCAGGGGACAGCAGCAGCCCTGAGTCTCTTCTGTTGGGAGTGGTCCGAGGT 1020  
 2014 GCAGTGGGCTAACAGGCTAGGCTTTGTTCTGTTGGGAGTGGTCCGAGGTCTCAGTCA 2073  
 1021 GCAGTGGGCTAACAGGCTAGGCTTTGTTCTGTTGGGAGTGGTCCGAGGTCTCAGTCA 1080  
 2074 TCCCTCCTAAGCAATGGGAGGCCCCACTAGCCAGTCTTCCAGGATCAACTGGGAG 2133  
 1081 TCCCTCCTAAGCAATGGGAGGCCCCACTAGCCAGTCTTCCAGGATCAACTGGGAG 1140  
 2134 GTTAGATGGGGCCAGGGTCCCAAGTACTGATGAGCCGAGCCAGGTTGAGCTTCTGG 2193  
 1141 GTTAGATGGGGCCAGGGTCCCAAGTACTGATGAGCCGAGCCAGGTTGAGCTTCTGG 1200  
 2194 TGTCCAGTCCGGATCCCACTGAGATCTCATGCTCTCAGATAGTGGGAGGAGTCTTT 2253  
 1201 TGTCCAGTCCGGATCCCACTGAGATCTCATGCTCTCAGATAGTGGGAGGAGTCTTT 1260  
 2254 TGTCCAGTCCGGATCCCACTGAGCCTCATGCTGCTGGTGGTGGTCTGCTGGGAA 2313  
 1261 TGTCCAGTCCGGATCCCACTGAGCCTCATGCTGCTGGTGGTGGTCTGCTGGGAA 1320  
 2314 AAGCTTTGGGGCTTCTGTTGTTAACACAGAGGAGGAGGAGGAGTGGGAGGAGTCTTT 2373  
 1321 AAGCTTTGGGGCTTCTGTTGTTAACACAGAGGAGGAGGAGGAGTGGGAGGAGTCTTT 1380  
 2374 TGTCCAGCCTCCCGTCTGGTGGGAGGAGGAGGAGTGGTGGTGGTGGTGGTGGTGGT 2433  
 1381 TGTCCAGCCTCCCGTCTGGTGGGAGGAGGAGGAGTGGTGGTGGTGGTGGTGGTGGT 1440

RESULT 2

US-08-232-463-14  
 Sequence 14, Application US/08232463  
 Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.  
 APPLICANT: SCHEIFLINGER, F.  
 APPLICANT: FALKNER, F. G.  
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 4800 Diagonal Road, Suite 500  
 CITY: Alexandria  
 STATE: VA  
 COUNTRY: USA  
 ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08232463  
 FILING DATE:

CLASSIFICATION:

CLASSIFICATION: 435  
 PRIOR APPLICATION NUMBER: US/07935319

FILING DATE:

FILING DATE: EP 91 114 300.6  
 FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300  
 TELEFAX: (703)683-4109  
 TELELEX: 899149

SEQUENCE CHARACTERISTICS:

SEQUENCE CHARACTERISTICS: 14:  
 LENGTH: 7218 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 CLONE: pTZgt-F1s

JOURNAL  
 COMMENT

Submitted (20-SEP-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 [WARNING] On Nov 2, 2001 this sequence was replaced by a newer version gi:16596673.  
 On Oct 13, 2000 this sequence version replaced gi:10765092.

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc/index.shtml  
 -----  
 Center project name: H\_NH0795C01  
 -----  
 Summary Statistics -----  
 Sequencing vector: M13; 100%  
 Chemistry: Dye-terminator Big Dye; 0% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 189029 bases at least Q40  
 Consensus quality: 197040 bases at least Q30  
 Insert size: 217000; agarose-fp  
 Insert size: 205278; sum-of-contigs  
 Quality coverage: 3.74 in Q20 bases; agarose-fp  
 Quality coverage: 3.99 in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently consists of 32 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. \* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. \*  
 \* 1 1933: contig of 1933 bp in length  
 \* 1934 2033: gap of unknown length  
 \* 2034 4225: contig of 2192 bp in length  
 \* 4226 4325: gap of unknown length  
 \* 4326 6989: contig of 2664 bp in length  
 \* 6990 7089: gap of unknown length  
 \* 7090 9679: contig of 2590 bp in length  
 \* 9680 9779: gap of unknown length  
 \* 9780 12591: contig of 2812 bp in length  
 \* 12592 12691: gap of unknown length  
 \* 12692 15210: contig of 2519 bp in length  
 \* 15211 18683: contig of unknown length  
 \* 18684 18783: contig of 3373 bp in length  
 \* 18784 23200: contig of unknown length  
 \* 23201 23301: gap of unknown length  
 \* 23302 27524: contig of 4224 bp in length  
 \* 27525 32318: contig of 4694 bp in length  
 \* 32319 32418: gap of unknown length  
 \* 32419 37205: contig of 4787 bp in length  
 \* 37206 37305: gap of unknown length  
 \* 37306 42143: contig of 4838 bp in length  
 \* 42144 42244: gap of unknown length  
 \* 42245 47623: contig of 5379 bp in length  
 \* 47624 47723: contig of unknown length  
 \* 47724 54427: gap of unknown length  
 \* 54428 60002: contig of 5576 bp in length  
 \* 60003 60102: gap of unknown length  
 \* 60103 66820: contig of 6718 bp in length  
 \* 66821 66921: gap of unknown length  
 \* 66922 74094: contig of 7174 bp in length  
 \* 74095 81658: contig of 7464 bp in length  
 \* 81659 87920: contig of unknown length  
 \* 87921 88020: contig of 1162 bp in length  
 \* 88021 88121: gap of unknown length

-----  
 \* NOTE: This is a 'working draft' sequence. It currently consists of 32 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. \* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. \*  
 \* 1 1933: contig of 1933 bp in length  
 \* 1934 2033: gap of unknown length  
 \* 2034 4225: contig of 2192 bp in length  
 \* 4226 4325: gap of unknown length  
 \* 4326 6989: contig of 2664 bp in length  
 \* 6990 7089: gap of unknown length  
 \* 7090 9679: contig of 2590 bp in length  
 \* 9680 9779: gap of unknown length  
 \* 9780 12591: contig of 2812 bp in length  
 \* 12592 12691: gap of unknown length  
 \* 12692 15210: contig of 2519 bp in length  
 \* 15211 18683: contig of unknown length  
 \* 18684 18783: contig of 3373 bp in length  
 \* 18784 23200: contig of unknown length  
 \* 23201 23301: gap of unknown length  
 \* 23302 27524: contig of 4224 bp in length  
 \* 27525 32318: contig of 4694 bp in length  
 \* 32319 32418: gap of unknown length  
 \* 32419 37205: contig of 4787 bp in length  
 \* 37206 37305: gap of unknown length  
 \* 37306 42143: contig of 4838 bp in length  
 \* 42144 42244: gap of unknown length  
 \* 42245 47623: contig of 5379 bp in length  
 \* 47624 47723: contig of unknown length  
 \* 47724 54427: gap of unknown length  
 \* 54428 60002: contig of 5576 bp in length  
 \* 60003 60102: gap of unknown length  
 \* 60103 66820: contig of 6718 bp in length  
 \* 66821 66921: gap of unknown length  
 \* 66922 74094: contig of 7174 bp in length  
 \* 74095 81658: contig of 7464 bp in length  
 \* 81659 87920: contig of unknown length  
 \* 87921 88020: contig of 1162 bp in length  
 \* 88021 88121: gap of unknown length

OM nucleic - nucleic search, using sw model  
 Run on: December 6, 2004, 14:03:15 ; Search time 627 Seconds  
 (without alignments)  
 3,920 Million cell updates/sec

Title: US-09-999-920-100  
 Perfect score: 2754  
 Sequence: 1 gccagagcagcctcagctt.....aaaataaaagatgcgcc 2754

Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 0.5  
 Searched: 3 seqs, 446265 residues  
 Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : \*txt:\*  
 1: /staff/overflow/sdavid-tmp/dec04/yu920/10198585.txt:\*  
 2: /staff/overflow/sdavid-tmp/dec04/yu920/10765092.txt:\*  
 3: /staff/overflow/sdavid-tmp/dec04/yu920/10800346.txt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	2604.1	94.6	208378	3 AC079988	ACCESSION:AC079988
C 2	1821.8	66.2	187722	2 AC079988	ACCESSION:AC079988
C 3	623.101	22.6	187722	2 AC079988	ACCESSION:AC079988
C 4	29.6	1.1	208378	3 AC079988	ACCESSION:AC079988
C 5	24.4	0.9	50165	1 AC079988	ACCESSION:AC079988
C 6	23.8	0.9	50165	1 AC079988	ACCESSION:AC079988

ALIGNMENTS

AC079988 208378 bp DNA linear HTG 14-OCT-2000  
 Homo sapiens chromosome 2 clone Rp11-795C1, WORKING DRAFT SEQUENCE,  
 32 unordered pieces.

ACCESSION AC079988  
 VERSION AC079988.3 GI:10800346  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 208378)

REFERENCE  
 AUTHORS Waterston,R.H.  
 TITLE The sequence of Homo sapiens clone  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 208378)  
 AUTHORS Waterston,R.H.  
 TITLE Direct Submission

Page 2 of 6

```

* 88021 94235: contig of 6215 bp in length
* 94335: gap of unknown length
* 102394: contig of 7959 bp in length
* 102394: gap of unknown length
* 111827: contig of 9433 bp in length
* 111828 111927: gap of unknown length
* 119373: contig of 7446 bp in length
* 119373: gap of unknown length
* 119473: contig of 11283 bp in length
* 130757 130856: gap of unknown length
* 130857 143566: contig of 12710 bp in length
* 143567 143666: gap of unknown length
* 143667 156273: contig of 12607 bp in length
* 156274 156373: gap of unknown length
* 156374 174943: contig of 18570 bp in length
* 174944 175043: gap of unknown length
* 175044 203094: contig of 28051 bp in length
* 203095 203194: gap of unknown length
* 203195 204665: contig of 1471 bp in length
* 204666 204765: gap of unknown length
* 204766 205864: contig of 1099 bp in length
* 205865 205964: gap of unknown length
* 205965 207143: contig of 1179 bp in length
* 207144 207243: gap of unknown length
* 207244 208378: contig of 1135 bp in length.
FEATURES
source
1..208378
location/qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-795C1"
1..12953
/notes="assembly_name:Contig10"
misc_feature
1934..2033
/estimated_length=unknown
gap
2034..4225
/notes="assembly_name:Contig11"
gap
4226..4325
/estimated_length=unknown
misc_feature
4326..6989
/notes="assembly_name:Contig12"
gap
6990..7089
/estimated_length=unknown
misc_feature
7090..9679
/notes="assembly_name:Contig13"
gap
9680..9779
/estimated_length=unknown
misc_feature
9780..12591
/notes="assembly_name:Contig14"
gap
12592..12691
/estimated_length=unknown
misc_feature
12692..15210
/notes="assembly_name:Contig15"
gap
15211..15310
/estimated_length=unknown
misc_feature
15311..18683
/notes="assembly_name:Contig16"
gap
18684..18783
/estimated_length=unknown
misc_feature
18784..23200
/notes="assembly_name:Contig17"
gap
23201..23300
/estimated_length=unknown
misc_feature
23301..27524
/notes="assembly_name:Contig18"
gap
27525..27624
/estimated_length=unknown
misc_feature
27625..32318
/notes="assembly_name:Contig19"
gap
32319..32418
/estimated_length=unknown
misc_feature
32419..37205

```

```

/notes="assembly_name:Contig20"
37206..37305
/estimated_length=unknown
misc_feature
37306..42143
/notes="assembly_name:Contig21"
42144..42243
/estimated_length=unknown
misc_feature
42244..47622
/notes="assembly_name:Contig22"
47623..47722
/estimated_length=unknown
misc_feature
47723..54326
/notes="assembly_name:Contig23"
54327..54426
/estimated_length=unknown
misc_feature
54427..60002
/notes="assembly_name:Contig24"
60003..60102
/estimated_length=unknown
misc_feature
60103..66820
/notes="assembly_name:Contig25"
66821..66920
/estimated_length=unknown
misc_feature
66921..74094
/notes="assembly_name:Contig26"
74095..74194
/estimated_length=unknown
misc_feature
74195..81659
/notes="assembly_name:Contig27"
81660..81758
/estimated_length=unknown
misc_feature
81759..87920
/notes="assembly_name:Contig28"
87921..88020
/estimated_length=unknown
misc_feature
88021..94235
/notes="assembly_name:Contig29"

```

```

Query Match          94.6%; Score 2604.1; DB 3; Length 208378;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 19; Indels 9; Gaps 8;

```

```

QY 1 GCCAAGCAGCCTCAGCTGGCAGGTTGGAGATGACTGCTGTCCTCCCTCGCATTTGG 60
Db 122926 GCCAAGCAGCCTCAGCTGGCAGGTTGGAGATGACTGCTGTCCTCCCTCGCATTTGG 122867
QY 61 GGAARACAGCTCCCTCGGTAGCTCGATGATCCCTTTTGTATCTTTGTGACCTCCCTGGA 120
Db 122866 GGAARACAGCTCCCTCGGTAGCTCGATGATCCCTTTTGTATCTTTGTGACCTCCCTGGA 122807
QY 121 GAGTGGATGACCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTCACCTGGCCGATGT 180
Db 122806 GAGTGGATGACCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTCACCTGGCCGATGT 122747
QY 181 CCCAGAGCAGGCAAGGTTCTAGAGCGGGTCTCCACATGACTGGCTTCCACACAGCA 240
Db 122746 CCCAGAGCAGGCAAGGTTCTAGAGCGGGTCTCCACATGACTGGCTTCCACACAGCA 122687
QY 241 CTTCCGCTGGGTTGCACTGCTGTTGTCATCTTACCGGTCAGGGTTGCAGGTAGGAAT 300
Db 122686 CTTCCGCTGGGTTGCACTGCTGTTGTCATCTTACCGGTCAGGGTTGCAGGTAGGAAT 122627
QY 301 GTTTGTACCTCTTCTGATTTGCCACTCTCTCCATGCTCCCTAGGACAGGCTTGGAG 360
Db 122626 GTTTGTACCTCTTCTGATTTGCCACTCTCTCCATGCTCCCTAGGACAGGCTTGGAG 122567
QY 361 GGCAGTGAAGGCTGGTTCAGGCACCCAGGCTCTCTTTGGACCTTCCCGGGGACCCCT 420
Db 122566 GGCAGTGAAGGCTGGTTCAGGCACCCAGGCTCTCTTTGGACCTTCCCGGGGACCCCT 122508
QY 421 GAGAGCTCTGAAAACCCCACTTAGCTTCCAGACCTTTCTCAAAGCTTCTCCCTGGCTT 480
Db 122507 GAGAGCTCTGAAAACCCCACTTAGCTTCCAGACCTTTCTCAAAGCTTCTCCCTGGCTT 122448

```





Page 4 of 8

Db 120235 TGIPTACTGTACCAATAGTGGCGAGAGAGTGTGTATGTGGGATCTGTGCTTGGGTTAGAA 120236

Qy 2701 TGCAAAATAAACTCACATTTGTAAGAAA 2728

Db 120235 TGCAAAATAAACTCACATTTGTAAGAAA 120208

AC079988/2 AC079988 187722 bp DNA linear HTG 11-OCT-2000
LOCUS AC079988/2 AC079988 187722 bp DNA linear HTG 11-OCT-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-795C1, \*\*\* SEQUENCING IN
PROGRESS \*\*\*, 57 unordered pieces.
ACCESSION AC079988 GI:10765092
VERSION AC079988.2
KEYWORDS HTG; HTGS PHASE1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 187722)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 187722)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT [WARNING] On Oct 13, 2000 this sequence was replaced by a newer
version gi:10800346.
On Oct 11, 2000 this sequence version replaced gi:10198585.

Center: Washington University Genome Sequencing Center
Web site: http://genome.wustl.edu/Gsc/index.shtml
----- Project Information -----
\* NOTE: This is a 'working draft' sequence. It currently
\* consists of 57 contigs. The true order of the pieces
\* is not known and their order in this sequence record is
\* arbitrary. Gaps between the contigs are represented as
\* runs of N, but the exact sizes of the gaps are unknown.
\* This record will be updated with the finished sequence
\* as soon as it is available and the accession number will
\* be preserved.
\* 1302: contig of 1302 bp in length
\* 1403: gap of unknown length
\* 2789: contig of 1387 bp in length
\* 2790: gap of unknown length
\* 4095: contig of 1205 bp in length
\* 4194: gap of unknown length
\* 5356: contig of 1162 bp in length
\* 5456: gap of unknown length
\* 5457: contig of 1212 bp in length
\* 6669: gap of unknown length
\* 6769: contig of 1099 bp in length
\* 7868: gap of unknown length
\* 7868: contig of 1497 bp in length
\* 9485: gap of unknown length
\* 9565: contig of 1512 bp in length
\* 11077: gap of unknown length
\* 11177: gap of unknown length
\* 12521: contig of 1344 bp in length
\* 12620: gap of unknown length
\* 12621: contig of 1432 bp in length
\* 14053: gap of unknown length
\* 14152: contig of 1723 bp in length
\* 15875: gap of unknown length
\* 15976: contig of 1415 bp in length
\* 17391: gap of unknown length
\* 17491: contig of 2079 bp in length
\* 19569: gap of unknown length
\* 19669: gap of unknown length

19570 20689: contig of 1020 bp in length
20690 20789: gap of unknown length
20790 22686: contig of 1897 bp in length
22687 22786: gap of unknown length
22787 25402: contig of 2616 bp in length
25403 25502: gap of unknown length
25503 27872: contig of 2370 bp in length
27873 27972: gap of unknown length
27973 29881: contig of 2009 bp in length
29882 30081: gap of unknown length
30082 32024: contig of 1943 bp in length
32025 32124: gap of unknown length
34083: contig of 1959 bp in length
34183: gap of unknown length
37046: contig of 2863 bp in length
37146: gap of unknown length
39347: contig of 2201 bp in length
39447: gap of unknown length
42510: contig of 3063 bp in length
42610: gap of unknown length
45752: contig of 3142 bp in length
45852: gap of unknown length
48396: gap of unknown length
50281: contig of 1885 bp in length
50381: gap of unknown length
52838: contig of 2457 bp in length
52939: gap of unknown length
55722: contig of 2784 bp in length
55822: gap of unknown length
58127: contig of 2305 bp in length
58227: gap of unknown length
60561: contig of 2634 bp in length
60961: gap of unknown length
64066: contig of 3105 bp in length
64166: gap of unknown length
67176: contig of 3010 bp in length
67276: gap of unknown length
69558: contig of 2282 bp in length
69658: gap of unknown length
72835: contig of 3177 bp in length
72935: gap of unknown length
75336: contig of 2601 bp in length
75636: gap of unknown length
79373: contig of 3737 bp in length
79473: gap of unknown length
82535: contig of 3062 bp in length
82635: gap of unknown length
85680: contig of 3045 bp in length
85780: gap of unknown length
89885: contig of 4105 bp in length
89985: gap of unknown length
94632: contig of 4547 bp in length
94732: gap of unknown length
98426: contig of 3794 bp in length
98526: gap of unknown length
102557: contig of 4031 bp in length
102657: gap of unknown length
108036: contig of 5579 bp in length
108136: gap of unknown length
113024: contig of 4888 bp in length
113124: gap of unknown length
116427: contig of 3303 bp in length
116527: gap of unknown length
120056: contig of 3529 bp in length
120156: gap of unknown length
120157: contig of 5011 bp in length
125267: gap of unknown length
13332: contig of 6065 bp in length
13432: gap of unknown length
136663: contig of 5231 bp in length
136763: gap of unknown length
141035: contig of 4272 bp in length